AFRICAN ELEPHANT CONSERVATION AND POPULATION GENETICS

BY

ADAM L. BRANDT

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Animal Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:

Associate Professor Alfred L. Roca, Chair
Teaching Associate Amy E. Fischer
Associate Professor Ripan S. Malhi
Professor Kelly S. Swanson

ABSTRACT

Despite advances in technology and management practices, countless species of wildlife continue to decline and become threatened with extinction, largely due to human activities such as poaching and habitat destruction. The field of conservation genetics aims to reduce the rate and severity of species decline by better understanding their genetics and making relevant information available to conservation and management entities. African elephants are in decline and the research here aims to answer important questions relevant to conservation efforts. (1) Phylogeographic patterns between nuclear and mitochondrial DNA in African elephants are often incongruent, which has been attributed to sex-biased dispersal and variance in reproductive success. To examine this, we sequenced the mitochondrial genome of two African forest elephants and examined the coalescent dates within the elephantid lineage. Comparing mitochondrial and nuclear coalescence dates, we found the ratio to be much greater than 0.25, which is consistent with the expectation that sex differences in dispersal and in variance of reproductive success would have increased the effective population size of mtDNA relative to nuclear markers in elephantids, thus contributing to the persistence of incongruent mtDNA phylogeographic patterns. (2) Past research on African elephant genetics has focused heavily on the phylogenetic relationship of forest and savanna elephants. Few studies have examined savanna elephant population genetics exclusively; those that have were limited in geographic scope or relied on mitochondrial DNA which has been shown to be a poor indicator of nuclear population structure. In this study we determined the extent of range wide, intra-species genetic variation for the African savanna elephant using multilocus genotype data. Our findings indicated that African savanna elephants have not undergone a population bottleneck within the last 2 to 4 N_E generations. Additionally, there was strong support for isolation by distance at the continental scale and there was evidence that localities in north-central Africa are distinct. (3) Many elephant populations in Africa are isolated within fragmented habitat and persist in low numbers. The elephants of Gash-Barka, Eritrea have become completely isolated, lacking any gene flow from other elephant populations. Using DNA isolated from dung, we examined nuclear and mitochondrial markers to better understand genetic variation and affinities to elephants elsewhere on the continent for conservation purposes. Elephants in Eritrea have low genetic diversity and a close affinity to savanna elephants in Eastern Africa. Conservation efforts

should aim to protect Eritrean elephants and their habitat in the short run, with restoration of habitat connectivity and genetic diversity as long-term goals. (4) While conservation genetics aims to reduce species decline, research is often limited by the availability of high quality DNA samples like blood or tissue. Obtaining these samples requires direct contact and handling that may be stressful and dangerous for the animals involved. Non-invasively collected samples such as feces have become increasingly preferred, but due to the presence of PCR inhibitors and non-target or fragmented DNA, the quality and quantity of information that can be obtained from these samples is severely limited. Advances in sequencing technology and enrichment techniques may prove useful in overcoming these limitations. We describe methods of capture hybridization to enrich low quality DNA samples, such as dung and ivory, for use with next-generation sequencing technology with African elephants serving as a model.

For Lily...

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Alfred Roca for giving me the opportunity to join his lab over a decade ago. That opportunity and the subsequent years of mentoring have helped me reach this point in my academic and professional career; for that I am extremely grateful and owe him a tremendous debt of gratitude. I would also like to thank my committee members, Drs. Amy Fischer, Ripan Malhi and Kelly Swanson for their advice, feedback and time that they spent helping me complete my doctoral degree.

I would like to thank my wife, Jess, for her insights, love and support. Few spouses are capable of discussing research projects, proof reading manuscripts and proposals, or assisting in experiments; but she helped with it all. I could not have reached this point without all she has done. I would also like to thank my parents and family for their support in my academic endeavors. It has been a long and difficult journey; I hope that I have made you all proud.

To all of my undergraduate research assistants, particularly Evan Emmel and McKenzie Canty, thank you for your hard work and dedication in the lab. Than you to my lab mates, Dr. Yasuko Ishida, Kai Zhao, Christina Ruiz and Tolulope Perrin-Stowe, for technical assistance, project discussions, and most important – fresh brewed coffee.

I would like to thank my co-authors and collaborators, Victor A. David, Nicholas J. Georgiadis, Yohannes Hagos, Jeheskel Shoshani, Sandra L. Shoshani, and Yohannes Yacob. Their assistance in manuscript preparation was invaluable. I thank V. Buckley-Beason, F. Hussain, J. Koontz, M. Malasky, and R. Hanson for technical or other assistance. For samples, I would like to thank Mike Chase and Kelly Landen from Elephants without Borders; Daniel Pouakouyou, Jean Louis Mamy, and Toupou Koighae from Fauna and Flora International; A. Turkalo and J. M. Fay, and A. Le´cu at the Parc Zoologique de Paris-Vincennes, France; the Elephant Research Foundation; and the Laboratory of Genomic Diversity. For permission to collect samples, I thank the Ministry of Agriculture, Government of Eritrea, in Asmara, Barentu and Haicota; the governments of Botswana, Cameroon, the Central African Republic, Democratic Republic of the Congo, Republic of the Congo, Gabon, Kenya, Namibia, South Africa, Tanzania, and Zimbabwe.

Generous funding for this research came from USFWS African Elephant Conservation Fund. Research facilities and support came from the Department of Animal Sciences, College of ACES, and the University of Illinois at Urbana-Champaign.

TABLE OF CONTENTS

CHAPTER 1: GENERAL INTRODUCTION

The order Proboscidea is represented by only three living species in the family Elephantidae, including the Asian elephant (*Elephas maximus*), African savanna elephant (*Loxodonta africana)*, and African forest elephant (*Loxodonta cyclotis*). Fossil records indicate that the most recent common ancestor of all elephantids diverged in the late Miocene and established three lineages (*Loxodonta*, *Elephas*, and *Mammuthus*) by 4 million years ago (Mya) (Maglio 1973). *Loxodonta* and *Elephas* are recognized as deeply divergent genera (Maglio 1973; Rohland *et al.* 2010), while the position of *Mammuthus* among elephantids had long been debated. Some studies using partial mitochondrial DNA sequences supported a shared ancestry between *Loxodonta* and *Mammuthus* (Barriel *et al.* 1999; Debruyne *et al.* 2003a); however, recent phylogenies generated utilizing full mitochondrial genomes (Krause *et al.* 2006; Rogaev *et al.* 2006; Rohland *et al.* 2007) and several hundred nuclear loci (Rohland *et al.* 2010) have supported a shared ancestor of *Mammuthus* and *Elephas*. Elephantid distribution has spanned much of the globe including North America, Europe, Asia and Africa (Maglio 1973). *Elephas* and *Loxodonta* lineages were both present in Africa until the late Pleistocene when *Elephas (recki) iolensis* went extinct leaving *Loxodonta* as the only elephantid in Africa (Maglio 1973; Sanders *et al.* 2010).

There are two extant species of *Loxodonta*, *L. africana* and *L. cyclotis* which both inhabit sub-Saharan Africa. The ancestry of *Loxodonta* has been somewhat uncertain, early studies have classified modern elephants as descendants of *L. adaurora* (Maglio 1973); however reanalysis and new fossils suggest that extant African elephants are directly descended from *L. exoptata* (Sanders *et al.* 2010). Taxonomic delineation of extant African elephants has been debated, with as many as 22 forms described as different species or subspecies that have since been redefined as being either forest or savanna elephants (Benjamin 2014- *In preparation*). African forest elephants occupy tropical rainforests while African savanna elephants are found in savanna bush and lightly forested regions (Grubb *et al.* 2000). African elephants may be named for the habitat in which they occur, but they do not occupy one habitat exclusively (Roth & Douglas-Hamilton 1991). Forest and savanna elephants can be distinguished from one another in the field by several features; savanna elephants have large triangular shaped ears, as well as forward and outward curved tusks, whereas forest elephants' ears are smaller and rounded, the tusks are thinner and

straighter, and the overall body shape is smaller and more compact (Grubb *et al.* 2000). Despite differences in physical appearances, the ranking of forest and savanna elephants as unique species or subspecies has been the subject of debate (Blanc *et al.* 2007; Debruyne 2005; Eggert *et al.* 2002; Groves & Grubb 2000; Johnson *et al.* 2007; Roca *et al.* 2001).

Early molecular studies of elephants found deeply divergent mitochondrial DNA (mtDNA) haplotypes within African elephants (Barriel *et al.* 1999; Georgiadis *et al.* 1994); the first published forest elephant mtDNA sequence revealed that African forest and savanna elephant divergence is as deep as the divergence between *Mammuthus* and *Elephas* (Barriel *et al.* 1999). Neither study made any recommendations about the ranking of African elephants, but their findings increased interest in the classification of African forest and savanna elephants. Morphometric measurements of 295 African elephant skulls were compared using canonical discriminant functions analyses (Groves & Grubb 2000), whereby Groves and Grubb (2000) concluded that African forest elephants and savanna elephants constitute two diagnosably distinct species. Molecular examinations using multiple nuclear genes supported this relationship, finding two reciprocally monophyletic clades corresponding to forest and savanna elephants (Comstock *et al.* 2002; Roca *et al.* 2001). Reports of hybridization with savanna elephants have been the main argument used to discredit classifying forest elephants as a distinct species (Barriel *et al.* 1999). Generally, reproductive isolation is the guiding criterion for species delineation and according to the biological species concept, "species are groups of interbreeding natural populations that are reproductively isolated from other such groups" (Mayr 1942). Roca and others (2001) found that the nuclear gene introgression between forest and savanna elephant populations was limited, thereby suggesting that the hybrid zone is fairly narrow, and "hybridization in nature is rare and perhaps minimized by behavioral or physiological reinforcement." Isolating mechanisms that favor homogeneous matings, physical or behavioral, support the attainment of species rank (Dobzhansnky 1941; Mayr 1942).

A comparison of phylogenies constructed using mtDNA and nuclear DNA revealed a discordant pattern between the two types of markers. Expanded datasets including elephants from west Africa (Eggert *et al.* 2002); (Johnson *et al.* 2007; Nyakaana *et al.* 2002) and those with known forest or savanna elephant morphology (Debruyne 2005) were able to reproduce the two previously described deeply divergent mtDNA clades (Barriel *et al.* 1999; Georgiadis *et al.* 1994). Yet, the observed clades were not reciprocally monophyletic, corresponding to a savanna

and forest elephant divergence as seen using nuclear DNA (Roca *et al.* 2001). Eggert and others (2002) analyzed microsatellite loci in addition to mtDNA and found the patterns to be "complex and defy attempts to divide the species into reciprocally monophyletic groups". They concluded that the phylogenetic pattern of African elephants can best be explained by a three taxa model: 1) forest elephants of central Africa, 2) forest and savanna elephants of west Africa and 3) savanna elephants of central east and southern Africa (Eggert *et al.* 2002). Johnson and others (2007) also concluded that the mtDNA patterns observed do not support a two species model and "the classification of species into savannah and forest may not reflect their evolutionary history but simply the habitat in which they currently exist" (Johnson *et al.* 2007). Debruyne (2005) designated the two clades "F" and "S", and found the S clade contained only individuals with savanna elephant morphology and the F clade contained individuals with all possible morphologies (savanna elephant, forest elephant and intermediate). He concluded that the paraphyletic relationship of forest and savanna elephants did not support any phylogenetic species concept for two or more species and attributed the differing results from previous studies to differential character sampling (Eggert *et al.* 2002) and limited sampling locations (Roca *et al.* 2001). Debruyne (2005) also noted that the persistence of F clade haplotypes nearly as frequent as S clade haplotypes in some savanna elephant populations conflicts with the assertion of a narrow hybrid zone (Roca *et al.* 2001) and that the extent of interbreeding violates the biological species concept (Mayr 1942) and thus species-level.

A comparison of mtDNA and sex-linked nuclear loci (Roca *et al.* 2005) revealed the same nuclear (Comstock *et al.* 2002; Roca *et al.* 2001) and mtDNA (Barriel *et al.* 1999; Debruyne 2005; Eggert *et al.* 2002; Georgiadis *et al.* 1994) phylogenetic patterns that had been observed previously. Roca and O'Brien (2005) suggested that the reciprocal monophyly among nuclear genes and the paraphyletic relationship in mtDNA is a result of sex differences in social structure and mating habits of African elephants resulting in unsuccessful reproduction by hybrid offspring. Female elephants are philopatric and remain with their natal herd (Archie *et al.* 2007); whereas, males are ejected from the herd upon sexual maturity and subsequently facilitate gene flow between herds (Archie *et al.* 2007; Hollister-Smith *et al.* 2007). As a result, nuclear genes are exchanged between elephant herds while mtDNA is bound to the herd. In areas where forest elephant and savanna elephant ranges overlap, smaller forest elephant males and hybrid males are unlikely to out-compete larger savanna elephant males for mating opportunities due to

intense male-male competition and increased mating success of older, larger males (Hollister-Smith *et al.* 2007; Poole 1989a; Poole 1989b). Low mating success for forest and hybrid male elephants would result in unidirectional gene flow of savanna elephant nuclear alleles into forest elephant herds. Paleoclimate data suggest that forest habitat coverage varied with glacial cycles during the last 2.58 million years (DeMenocal 2004; Gibbard & Head 2010; Maley 2001) and likely extended into geographic regions that currently feature savanna habitat. The presence of F clade mtDNA in savanna elephant populations deep in savanna habitat (Debruyne 2005; Roca *et al.* 2005) can be explained by hybridization between forest and savanna elephants when forest habitat boundaries receded, leaving isolated forest elephant herds. Given the disproportionate rate of reproductive success of larger male elephants (Hollister-Smith *et al.* 2007), savanna male elephants would be expected to out-compete forest and hybrid male elephants, resulting in repeated backcrossing between hybrid female elephants and savanna male elephants (Roca *et al.* 2005; Roca *et al.* 2007). If this were not the case, and if hybridization was wide-spread, supporting a one species model as suggested by Debruyne (2005), then 1) savanna locations would have a high correlation between mtDNA and elephant morphology, 2) forest elephant and intermediate morphologies would be common in savanna populations where forest elephant mtDNA predominates, 3) intermediate morphologies would be common across forest and savanna habitats, and 4) distribution of elephants with intermediate morphologies would be common across the savanna wherever forest and savanna elephant mtDNA haplotypes are both found (Roca *et al.* 2007). Presumably under a single species model with a wide hybrid zone, nuclear genotypes and mtDNA haplotypes would have a linear relationship whereby the proportion of forest elephant and savanna elephant nuclear genotypes would reflect the proportion of F and S clade mtDNA haplotypes, respectively, in a population (Ishida *et al.* 2011b). Elephants in savanna locations possess exclusively savanna genotypes regardless of mtDNA haplotype and S clade mtDNA haplotypes are absent in forest populations (Ishida *et al.* 2011b). The higher rate of mtDNA introgression relative to nuclear DNA is evident in numerous other species with male-biased dispersal (Petit & Excoffier 2009). Furthermore, elephants are not the only species exhibiting mtDNA introgression; several studies have observed mitochondrial introgression between well-established species including chipmunks (Good *et al.* 2008), Columbian mammoth (Enk *et al.* 2011), and polar bear (Hailer *et al.* 2012).

The taxonomic standing of African elephants has been thoroughly investigated; however, the International Union for Conservation of Nature and Natural Resources (IUCN) only recognizes one species of African elephant for fear of leaving hybrid elephants with an uncertain taxonomic and conservation status (Blanc *et al.* 2007). African elephants were listed as threatened under the Endangered Species Act in 1978 and were awarded further protection in 1989 when they were added to Appendix I of the Convention on International Trade in Endangered Species of Wild Flora and Fauna (CITES), thus prohibiting commercial trade of elephant products; yet, poaching and habitat loss continue to be a major threat to elephant conservation (Barnes 1999; Douglas-Hamilton 1987; Wasser *et al.* 2009; Wittemyer *et al.* 2014). Examining phylogenetic patterns, population genetic structure and exploring new ways to utilize alternative DNA sources will aid in the conservation and management of elephants. To expand the knowledge base of African elephant conservation genetics, three questions have been selected for further research:

(1)The phylogenetic patterns of nuclear and mitochondrial DNA in African elephants are incongruent, which has been attributed to sex differences in dispersal and variance of reproductive success (Ishida *et al.* 2011b; Roca *et al.* 2005; Roca *et al.* 2007). If this is true, then lower dispersal and lower variance in reproductive success among females would increase mtDNA effective population size relative to nuclear DNA which would be evident as a ratio of coalescence dates greater than 0.25 (value expected in populations without differences in dispersal or reproductive success). To examine this relationship, the complete mtDNA genome of African forest elephants was sequenced to determine mtDNA coalescence relative to other elephantids.

(2) Numerous studies have been conducted on African elephant genetics, examining the phylogenetic relationship between forest and savanna elephants (Debruyne 2005; Eggert *et al.* 2002; Johnson *et al.* 2007; Roca *et al.* 2005; Roca *et al.* 2001). Few studies have examined the population genetic structure of African elephants, and those that did, focused on regional populations (Nyakaana & Arctander 1999; Okello *et al.* 2008) or relied heavily on mtDNA (Georgiadis *et al.* 1994; Nyakaana *et al.* 2002), which has been shown to be a poor indicator of population structure among African elephants (Ishida *et al.* 2011b; Roca *et al.* 2005; Roca *et al.* 2007). None of these studies has specifically examined the intra-species relationship among African savanna elephants. Previous studies have suggested that African savanna elephants have

low genetic diversity due to a founder effect when *Elephas* went extinct in Africa (Roca *et al.* 2001) or due to a recent continent-wide population expansion within a short evolutionary period (Georgiadis *et al.* 1994). Using multilocus genotype data, African savanna elephants were examined for genetic diversity, gene flow, changes in effective population size, population and geographic structure.

(3) Numerous elephant populations persist in isolated and fragmented habitat. Elephants have historically occurred throughout Eritrea, though they were believed to have been extirpated by the early 20th century (Gowers 1948). Recent surveys described a small population on protected land in the southwestern region of Gash-Barka. The elephants in Eritrea are thought to be isolated from other populations (Shoshani *et al.* 2000; Shoshani *et al.* 2004). Low genetic diversity and severe inbreeding is a major concern for the preservation of this population; a better understanding of their genetics will allow for sound management decisions. To accomplish this, DNA was isolated from non-invasively collected dung samples. Diagnostic, mitochondrial and nuclear DNA markers were amplified and analyzed to examine genetic diversity and ancestry.

CHAPTER 2: FOREST ELEPHANT MITOCHONDRIAL GENOMES REVEAL THAT ELEPHANTID DIVERSIFICATION IN AFRICA TRACKED CLIMATE TRANSITIONS

Abstract

Among elephants, the phylogeographic patterns of mitochondrial (mt) and nuclear markers are often incongruent. One hypothesis attributes this to sex differences in dispersal and in the variance of reproductive success. We tested this hypothesis by examining the coalescent dates of genetic markers within elephantid lineages, predicting that lower dispersal and lower variance in reproductive success among females would have increased mtDNA relative to nuclear coalescent dates. We sequenced the mitochondrial genomes of two forest elephants, aligning them to mitogenomes of African savanna and Asian elephants, and of woolly mammoths, including the most divergent mitogenomes within each lineage. Using fossil calibrations, the divergence between African elephant F and S clade mitochondrial genomes (originating in forest and savanna elephant lineages, respectively) was estimated as 5.5 million years ago (Mya). We estimated that the (African) ancestor of the mammoth and Asian elephant lineages diverged 6.0 Mya, indicating that four elephantid lineages had differentiated in Africa by the Miocene-Pliocene transition, concurrent with drier climates. The coalescent date for forest elephant mtDNAs was ca. 2.4 Mya, suggesting that the decrease in tropical forest cover during the Pleistocene isolated distinct African forest elephant lineages. For all elephantid lineages, the ratio of mtDNA to nuclear coalescent dates was much greater than 0.25. This is consistent with the expectation that sex differences in dispersal and in variance of reproductive success would have increased the effective population size of mtDNA relative to nuclear markers in elephantids, contributing to the persistence of incongruent mtDNA phylogeographic patterns.

Introduction

A number of studies have indicated that the African savanna or bush elephant (*Loxodonta africana*) is a distinct species from the African forest elephant (*L. cyclotis*) (Comstock *et al.* 2002; Groves & Grubb 2000; Grubb *et al.* 2000; Roca *et al.* 2005; Roca *et al.* 2001; Rohland *et al.* 2010). Analyses of DNA sequences have demonstrated that little or no nuclear gene flow occurs between forest elephant and savanna elephant populations, providing evidence that the two groups would satisfy the definition, under the biological species concept, that "species are groups of interbreeding natural populations that are reproductively isolated from other such groups" (Mayr 1963; Meier & Wheeler 2000; Petit & Excoffier 2009; Roca *et al.* 2005; Roca *et al.* 2007; Roca *et al.* 2001). Corroborating this evidence are reports that forest and savanna elephants can be completely distinguished morphologically using a discriminant function (Groves & Grubb 2000), that between forest and savanna elephants Fst values greater than 0.90 have been estimated for nuclear sequences and microsatellites (Comstock *et al.* 2002; Roca *et al.* 2001), that phylogenies inferred using nuclear markers place forest and savanna elephants into reciprocally monophyletic clades (Comstock *et al.* 2002; Roca *et al.* 2001), and that the divergence between them has been estimated as 2.6 to 5.6 Mya using nuclear markers (Rohland *et al.* 2010).

Mitochondrial (mt) DNA phylogeographic patterns among elephantids are often discordant with those of nuclear DNA markers or morphology (Debruyne 2005; Enk *et al.* 2011; Ishida *et al.* 2011b; Lei *et al.* 2012; Roca *et al.* 2005). Among African elephants, mtDNA forms two deeply distinctive clades, designated clades F and S (Debruyne 2005), originating in forest and savanna elephants, respectively (Ishida *et al.* 2011b). Forest elephants do not carry S clade mtDNA, while some nuclear alleles that are common or fixed across savanna elephants are not found in forest elephants (Ishida *et al.* 2011b; Roca *et al.* 2005), suggesting that forest elephant populations may be completely or almost completely isolated from gene flow from savanna elephant populations. By contrast, many African savanna elephant individuals and populations are known to carry F clade mtDNA (Debruyne 2005; Lei *et al.* 2008; Roca *et al.* 2005), although almost all of these show no evidence of carrying forest elephant nuclear alleles (Lei *et al.* 2009; Roca *et al.* 2005). This mito-nuclear incongruence has suggested that species isolation mechanisms exist that largely prevent hybrid males, but not hybrid females, from reproducing

successfully among savanna elephants (Roca *et al.* 2005; Roca *et al.* 2007; Roca & O'Brien 2005). Nonetheless, based on both simulations and empirical surveys of interspecies gene flow patterns (Currat *et al.* 2008; Petit & Excoffier 2009), male-mediated dispersal in savanna elephants (Archie *et al.* 2008) would be expected to facilitate the widespread introgression of mtDNA but not nuclear alleles from forest to savanna elephant populations.

Among Asian elephants (*Elephas maximus*), two deeply divergent mtDNA clades separated 1.6 to 2.1 Mya. These mtDNA clades do not appear to correspond to recognized subspecies subdivisions (Deraniyagala 1955), or to nuclear genetic subdivisions in Asian elephants (Lei *et al.* 2012), but may instead reflect the impact of Pleistocene glacial cycles on the climate and habitats of Asia (Fernando *et al.* 2000; Fernando *et al.* 2003; Fleischer *et al.* 2001; Vidya *et al.* 2009). In an extinct elephantid, the woolly mammoth (*Mammuthus primigenius*), two clades with ca. 1 to 2 Mya divergence have been detected (Gilbert *et al.* 2008). One clade appears to be more geographically restricted than the other, and to have gone extinct earlier (Gilbert *et al.* 2008). The two woolly mammoth mtDNA clades, and further subdivisions detected within them, have provided the framework for hypotheses involving the potential effects of speciation events, migrations, habitat changes, and glacial cycles on woolly mammoths (Barnes *et al.* 2007; Debruyne *et al.* 2008; Gilbert *et al.* 2008; Miller *et al.* 2008). Recently, woolly mammoth mtDNA was found to have introgressed into a morphologically distinctive species, the Columbian mammoth (*Mammuthus columbi*) (Enk *et al.* 2011).

Sex differences in dispersal and in the variance of reproductive success have been well documented in living elephantids (Archie *et al.* 2007; Hollister-Smith *et al.* 2007; Poole *et al.* 2011; Roca *et al.* 2007). Elephant females reaching maturity remain with their natal core social group or "herd" (Archie *et al.* 2007; Hollister-Smith *et al.* 2007) and females do not typically migrate between herds (Archie *et al.* 2007; Hollister-Smith *et al.* 2007), thus the mitochondrial genome is necessarily coupled to the geographic range of the core social group. By contrast, males leave the natal herd and mediate gene flow between herds and across the landscape (Archie *et al.* 2007; Hollister-Smith *et al.* 2007; Poole *et al.* 2011; Roca *et al.* 2007). Furthermore, male-male competition is intense, and in any generation only a small proportion of males are reproductively successful (Archie *et al.* 2007; Hollister-Smith *et al.* 2007; Rasmussen *et al.* 2008). We have previously formulated the hypothesis that these well-documented sex

differences in elephants have enabled the observed phylogeographic incongruence in mito-nuclear patterns (Ishida *et al.* 2011b).

Since lower female dispersal and higher variance in male reproductive success would have increased mtDNA effective population sizes relative to those of nuclear loci (Hedrick 2007; Hoelzer 1997), then if the matrilocal and matrilineal social structure and intense male reproductive competition are responsible for the observed mito-nuclear incongruence (Ishida *et al.* 2011b), one would predict that the ratio of coalescent dates for mtDNA relative to those of nuclear loci would have a value greater than the 0.25 expected in populations without sex differences in dispersal or in variance in reproductive success. Here, we examine the coalescent dates of mtDNA and nuclear loci in elephantids to test our expectation that sex differences in dispersal and in the variance of reproductive success have led to disproportionately high mtDNA effective population sizes and coalescent dates relative to those of nuclear markers. We generated full mitochondrial genome sequences for *Loxodonta cyclotis*, sequencing two forest elephants that carried haplotypes corresponding to the two most divergent mtDNA subclades within the species (Debruyne 2005; Roca *et al.* 2005). We aligned them to mitogenomes representing the most divergent mtDNA clades within each of the other elephantid lineages. We examined the relationships of forest elephants to other elephantid taxa, and estimated divergence and coalescent dates using fossil calibrations. We contrasted the ratio of mtDNA to nuclear coalescent dates within and across elephantid taxa. Finally, we compared our date estimates to the timing of climate transitions, finding that elephantid evolution in Africa may have been driven by drier climates at the end of the Miocene and by the initiation of glacial cycles at the start of the Pleistocene.

Materials and Methods

Elephant Samples and DNA Extraction

The study was conducted in compliance with the University of Illinois Institutional Animal Care and Use Committed (IACUC) approved protocol number 09036. Samples were obtained in full compliance with required CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora) and other permits. Two forest elephant DNA

samples were used. One was extracted from tissue collected using a biopsy dart from a female forest elephant (designated Lcy-DS1534 or DS1534) in the Dzanga Sangha Forest Reserve of the Central African Republic (Georgiadis *et al.* 1994). The other was extracted from a blood sample of "Coco," a male forest elephant (*Loxodonta cyclotis*) from Sierra Leone, designated Lcy-SL0001 or SL0001, generously provided by the Paris Zoo (Parc Zoologique de Paris-Vincennes, France). Nuclear DNA sequences had previously established that both individuals were forest elephants (Capelli *et al.* 2006; Ishida *et al.* 2011b; Roca *et al.* 2005; Roca *et al.* 2001); mtDNA sequences had established that SL0001 was from a sub-clade of mtDNA found in West Africa (Barriel *et al.* 1999; Debruyne 2005; Debruyne *et al.* 2003b; Eggert *et al.* 2002). DNA was extracted using a kit from Qiagen (SL0001), or a standard phenol-chloroform method (DS1534) (Sambrook *et al.* 1989).

Oligonucleotide Primers

The mitochondrial genome was amplified in eight long overlapping fragments. We designed eight PCR primer pairs (Table 2.1) using DNA sequences identified as being conserved (Murphy & O'Brien 2007) between the mtDNA genomes of the elephant (Hauf *et al.* 1999) and the aardvark (Arnason *et al.* 1999), or as being conserved across published elephantid sequences. In addition to the PCR primers, within each of the amplified regions we designed a set of additional sequencing primers to produce forward or reverse sequences within each of the eight amplified segments (Table 2.1); together these sets generated overlapping sequences spanning the entire mitochondrial genome, except for a repetitive non-coding region, containing a variable number of tandem repeats (VNTR), that has also not been reported for most previously generated proboscidean mitogenomes (Gilbert *et al.* 2008; Gilbert *et al.* 2007; Rogaev *et al.* 2006).

PCR and Sequencing

Genomic DNA (ca. 50 ng) underwent amplification by PCR in a 25 ul reaction volume containing 1x GeneAmp PCR buffer II, 1.5 mM $MgCl_2$, 0.2 mM each of the four deoxyribonucleoside 5'-triphosphates (dATP, dCTP, dGTP, and dTTP), 0.04 unit/ul AmpliTaq-GOLD DNA Polymerase (Applied Biosystems Inc. [ABI]), and 0.4 uM final concentration of each oligonucleotide primer. PCR was run with an initial step of 95°C for 9:45 min; with cycles of 20 sec at 94°C; followed by 30 sec at 60°C (first 3 cycles), 58°C (next 5 cycles), 56°C (5

cycles), 54°C (5 cycles), 52°C (5 cycles), or 50°C (final 22 cycles); followed by 3 min extension at 72°C; with a final extension after the last cycle of 7 min at 72°C. PCR products were enzyme-purified (Hanke & Wink 1994) and sequenced using the BigDye Terminator system (ABI). Extension products were purified using Sephadex G-50 (Amersham) and resolved on an ABI 3700 DNA Sequencer, or on ABI 3730xl capillary systems. The software *Sequencher* (Gene Codes Corporation) was used to trim and concatenate sequences and confirm open reading frames. DNA sequences have been deposited in GenBank (accession numbers: JN673263 and JN673264).

Avoidance of Nuclear DNA Sequences of Mitochondrial Origin (numts)

Steps were taken to minimize the possibility of amplifying numts. Primers were designed for conserved regions (see above), which would minimize the risk of mismatches in the target mtDNA versus the primers. Amplicons were several thousand bases in length, which would avoid those numts that were shorter. Obvious indicators of numts were not observed: PCR did not produce multiple bands; there were no sites that appeared heteroplasmic with secondary peaks; open reading frames were present in all coding regions (Figure 2.1)--these would have been disrupted in many numts; sequences overlapping between amplicons were identical in the region of overlap, minimizing the possibility that primer mismatches had led to selective amplification of numts over cytoplasmic mtDNA.

Sequence Alignments

Three mtDNA genome sequences have been generated for S clade African savanna elephants (Hauf *et al.* 1999; Murata *et al.* 2009; Rogaev *et al.* 2006), along with three for Asian elephants (Arnason *et al.* 2008; Maikaew *et al.* 2007; Rogaev *et al.* 2006), nineteen for woolly mammoths (Gilbert *et al.* 2008; Gilbert *et al.* 2007; Krause *et al.* 2006; Rogaev *et al.* 2006), and one Columbian mammoth mitogenome (Enk *et al.* 2011). The mtDNA genome has also been sequenced for one proboscidean that is outside of the family Elephantidae, the extinct American mastodon (*Mammut americanum*), which diverged from the lineage leading to elephantids some 24-30 Mya (Rasmussen & Gutierrez 2009; Rohland *et al.* 2007; Rohland *et al.* 2010; Sanders *et al.* 2010). We retrieved complete mtDNA genomic sequences from GenBank for seven elephantid individuals, including three African savanna elephants (GenBank: DQ316069,

NC000934, and AB443879) (Hauf *et al.* 1999; Murata *et al.* 2009; Rogaev *et al.* 2006), two Asian elephants (GenBank: NC005129.2 and AJ428946) (Arnason *et al.* 2008; Rogaev *et al.* 2006) and two woolly mammoths (GenBank: NC007596.2 and EU153453) (Gilbert *et al.* 2008; Krause *et al.* 2006). For some analyses, the mitogenomic sequence of an American mastodon (GenBank: NC_009574) (Rohland *et al.* 2007) was also included. Published genomes were aligned with the two newly generated African forest elephant sequences using the software CLUSTALW 2.0 (Larkin *et al.* 2007) in EBI Web Services (McWilliam *et al.* 2009); alignment output was visually inspected. The control region was excluded from analyses since it is subject to greater saturation and is less reliable as a molecular clock relative than other mtDNA regions (Ingman *et al.* 2000). Mismatches between the mtDNA genomes were visualized using the software Geneious (Drummond *et al.* 2010).

Phylogenetic Analyses

The Akaike Information Criterion (AIC) (Akaike 1974) was implemented using the software Modeltest 3.06 (Posada & Crandall 1998) to estimate the model of DNA sequence evolution that best fit the data. This was found to be the Tamura-Nei model with invariant site heterogeneity (TrN+I), with the following parameters: Base (base frequencies for A, C and G, with T inferred) = (0.3286 0.2523 0.1358); Nst (number of substitution types listed in a rate matrix) = 6 (the number of unique substitution types for this model is 3; PAUP* specifies this using a matrix with six values, four of which are identical); Rmat (rate matrix) = (1.0000 38.7425 1.0000 1.0000 55.3986 1.0000); Rates (distribution of rates at variable sites) = equal; and Pinvar (proportion of invariant sites) = 0.7045. These parameter values were used in PAUP*4.0b10 (Swofford 2002) for Neighbor Joining (NJ), minimum evolution (ME), and maximum likelihood (ML) phylogenetic methods. PAUP* was also used to infer a maximum parsimony (MP) tree. For all phylogenetic methods, exhaustive searches were conducted. Bootstrap resampling support was based on at least 100 replicates, in a full heuristic search with starting trees obtained by random stepwise addition and tree bisection-reconnection branch swapping for MP, ME and ML analyses. Bayesian phylogenetic inference was performed using BEAST v1.5.4 software (Drummond & Rambaut 2007), using the TN93 + I model of nucleotide substitution, determined using the AIC (Akaike 1974).

The full mitochondrial genome was partitioned into protein coding genes, rRNA+tRNA genes, H-strand genes and L-strand genes. The Akaike Information Criterion (AIC) (Akaike 1974) was implemented using the software Modeltest 3.06 (Posada & Crandall 1998) to estimate the model of DNA sequence evolution that best fit each partition. AIC parameter values were used in PAUP*4.0b10 (Swofford 2002) for maximum likelihood (ML) phylogenetic analysis, with each partition run individually. Although models cannot be applied to the maximum parsimony tree, PAUP* was also used to infer a maximum parsimony (MP) tree for each of the datasets. For both ML and MP methods, exhaustive searches were conducted and yielded topologies consistent with un-partitioned data; in each case relationships among the individuals on the tree were the same. Model parameters and tree statistics are as follows:

The model of evolution determined by Modeltest is listed here for each partition as follows: "Base" indicates the base frequencies for A, C and G, with T inferred (bases are "equal" when the model assumes frequencies are the same). "Nst" lists the number of substitution types listed in a rate matrix; the number of unique types may be inferred.

This is followed by "Rmat", the rate matrix. "Rates" indicates the distribution of rates at variable sites. "Pinvar" indicates the proportion of invariant sites:

For Protein coding genes: Tamura-Nei model with invariant site heterogeneity (TrN+I). Base=(0.3278 0.2544 0.1307) Nst=6 (unique Nst=3) Rmat=(1.0000 39.8690 1.0000 1.0000 54.9781 1.0000) Rates=equal Pinvar=0.6883.

For rRNA and tRNA genes: Tamura-Nei model with invariant site heterogeneity (TrN+I). Base=(0.3572 0.2134 0.1493) Nst=6 (unique Nst=3) Rmat=(1.0000 25.9384 1.0000 1.0000 56.4363 1.0000) Rates=equal Pinvar=0.8094

For H-strand: Tamura-Nei model with invariant site heterogeneity (TrN+I). Base=(0.3271 0.2495 0.1336) Nst=6 (unique Nst=3) Rmat=(1.0000 38.1881 1.0000 1.0000 53.5181 1.0000) Rates=equal Pinvar=0.6923

For L-strand: Tamura-Nei model with invariant site heterogeneity (TrN+I). Base=(0.3738 0.2753 0.1142) Nst=6 (unique Nst=3) Rmat=(1.0000 48.6352 1.0000 1.0000 89.7056 1.0000) Rates=equal Pinvar=0.7637

The AIC parameter values were also used in BEAST (Drummond & Rambaut 2007) for Bayesian posterior probability estimation, using a dataset consisting of the entire mitogenome (excluding control region), but with each partition allowed to follow a different model of

evolution. Additional settings included the use of an uncorrelated lognormal relaxed molecular clock model and a randomly generated starting tree; all other parameters remained as default settings. Posterior distributions were obtained by Markov chain Monte Carlo (MCMC) sampling from a total of 10,000,000 steps, with a discarded burn-in of 1,000,000; samples were drawn every 1,000 MCMC steps. Acceptable mixing and convergence to the stationary distribution were verified by inspection and plotting of posterior samples. Effective sample size values were above 200 for parameters.

Molecular Dating

Nodal divergence dates within the tree were estimated using the software BEAST (Drummond & Rambaut 2007), with fossil estimates used as prior dates for calibration. We utilized an uncorrelated lognormal relaxed molecular clock model, which permits the rate of molecular substitution to be uncorrelated across the tree while incorporating uncertainty in both tree topology and multiple fossil calibrations (Drummond *et al.* 2006). Fossil calibration priors were incorporated into the clock model as a normal distribution representing soft bounds (5^{th} and 95^{th} percentile corresponding to the minimum and maximum dates respectively), since maximum fossil calibration dates incorporated uncertainty in the fossil record (see below) (Forest 2009). We used a randomly generated starting tree. Posterior distributions were obtained by Markov chain Monte Carlo (MCMC) sampling from a total of 10,000,000 steps, with a discarded burn-in of 1,000,000; samples were drawn every 1,000 MCMC steps. Acceptable mixing and convergence to the stationary distribution were verified by inspection and plotting of posterior samples. Effective sample size values were above 200 for all parameters.

Two sets of prior date calibrations were used. The fossil calibration dates listed below were applied to one dataset that included only elephantid sequences and calibration dates, and then applied to a second dataset that included the American mastodon mitogenome as outgroup, and also used a mastodon-elephantid calibration in addition to the elephantid calibration priors. Thus we performed a total of four analyses, two with only the elephantid mitogenomes and two with the mastodon mitogenome included as outgroup. The two sets of prior dates used as calibrations were as follows:

(i) Our "broad" fossil calibration used the same fossil date range estimates as Rohland et al. (2010), except that no prior date estimates were used for the divergence between African

elephant F and S mtDNA clades. Rohland et al. (2010) set the date range for the mastodon-elephantid divergence as 24-30 Mya. The minimum was based on Sanders et al. (2010), who mention the Rasmussen and Gutierrez (2009) report of fauna from the Eragaleit Beds of Lothidok as including the earliest definitively known mammutid, dated to between >24-27 Mya (Rasmussen & Gutierrez 2009). The maximum was also based on Sanders et al. (2010), who mention a specimen from Chilga, Ethiopia (28-27 Ma) described as having mammutid features superimposed on an otherwise palaeomastodont-type molar. To this and to other estimates of maximum dates was added 2 My, to account for difficulty in recognizing differences during an initial period of divergence, and to account for inadequate fossil sampling (Rohland *et al.* 2010). The 2 My also approximates the maximum (1.8 My) of averages for elephantid species duration that had been calculated by Maglio using six intervals of one million years of time, with 3 to 14 elephantid species present within each interval (Maglio 1973).

The range for the split between *Loxodonta* and Eurasian elephantids was set as 4.2-9 Mya (we use "Eurasian" to refer to the clade including Asian elephants and woolly mammoths, while recognizing that both *Elephas* and *Mammuthus* originated in Africa, and the woolly mammoth range included North America) (Rohland *et al.* 2010). The minimum of 4.2 Mya was based on Sanders et al. (2010), who list *Elephas ekorensis*, as "the most ancient unequivocal representative of the genus *Elephas*" (Sanders *et al.* 2010). Among locations listed for this species, the oldest is 5.0-4.2 Mya (Sanders *et al.* 2010). This minimum also assumes that *Loxodonta* is paraphyletic. The maximum of 9 Mya does not assume that *Loxodonta* is paraphyletic, and is based on the older bound (7 Mya) of the age estimate of 6-7 Mya for the oldest *Loxodonta* fossil (Sanders *et al.* 2010; Vignaud *et al.* 2002), to which was added 2 Mya to the maximum date for the reasons noted above (Rohland *et al.* 2010).

The range for the split between *Elephas* and *Mammuthus* was set at 3-8.5 Mya (Rohland *et al.* 2010). Although the earliest reported fossil mammoth is *Mammuthus subplanifrons*, Sanders et al. (2010) state that "in the absence of associated crania, however, there is no certainty that this species is a mammoth", and that *M. africanavus* is "the earliest unambiguous evidence of the genus *Mammuthus* in Africa". In the associated table in Sanders et al. (2010), the latter species is listed as mid-late Pliocene, with dated locales at 3.5-3.0 Mya (Sanders *et al.* 2010). Rohland et al. therefore took 3 Mya as the minimum date for the *Elephas-Mammuthus* divergence, which also assumes that *Elephas* is paraphyletic (Rohland *et al.* 2010). The

maximum date for the split did not assume paraphyly for the taxa. For the questionable *M. subplanifrons*, the oldest date listed by Sanders et al. (2010) for any date range at a locale is 6 Mya, but there is an older "*Elephas nawataensis*" that dates to 4.2-6.5 Mya (Leakey & Harris 2003). Although Sanders et al. (2010) suggest that this taxon may be mis-assigned to the genus *Elephas* (Rohland *et al.* 2010), it was nonetheless included by Rohland et al. (2010) as the oldest potential *Elephas* fossil identified in the literature; adding 2 Mya to the maximum date for the reasons noted above provided the 8.5 Mya maximum (Rohland *et al.* 2010).

(ii) Our "narrow" fossil calibration range reduced the range of fossil date estimates in two ways. First we removed from consideration *Elephas nawataensis* and *Mammuthus subplanifrons* since Sanders et al. (2010) considered them to be potentially mis-assigned to their respective genera. Second, we considered all genera to be monophyletic, consistent with several fossil phylogenies (Maglio 1973; Shoshani & Tassy 1996). This made the minimum date estimate for the *Loxodonta*-Eurasian split equal to the lower bound of the age attributed to the oldest *Loxodonta* fossil, or 6 Mya (Sanders *et al.* 2010; Vignaud *et al.* 2002), while the maximum date for this divergence remained the same as for the "broad" fossil calibration. For the divergence between *Elephas* and *Mammuthus*, the oldest fossil recognized by Sanders et al. (2010) as definitely assigned to either genus is *Elephas ekorensis*. Among locations listed for this species, the oldest is 4.2-5.0 Mya (Sanders *et al.* 2010). Thus the maximum date for the calibration of the split between *Elephas* and *Mammuthus* is 7.0 Mya (including two million years added for reasons noted above). The assumption of monophyly would mean that this is also the fossil used to set a date minimum for the *Elephas*-*Mammuthus* split, of 4.2 Mya (Sanders *et al.* 2010). The "narrow" fossil calibration uses the same calibration dates as the "broad" fossil calibration for the mastodon-elephantid split (for the dataset that included a mastodon mitogenome) (Rohland *et al.* 2010). Like the "broad" fossil calibration, it used no prior estimate for the divergence between F and S African elephant mtDNA clades.

Results

Forest Elephant Mitochondrial Genomes

For the two forest elephants, we sequenced all 13 coding genes of the mtDNA genome, along with the ribosomal RNAs and transfer RNAs, and part of the control region. We sequenced 16,028 bp of mtDNA for forest elephant DS1534 and 16,156 bp of mtDNA for forest elephant SL0001. We generated an alignment of the two forest elephant sequences with the mitochondrial genomes of other elephantids, including all three previously sequenced mitogenomes for S clade savanna elephants. We also aligned two Asian elephant and two woolly mammoth mtDNA genomic sequences representing the deepest subdivisions present across the mtDNAs of those species (Fernando *et al.* 2000; Fleischer *et al.* 2001; Gilbert *et al.* 2008; Krause *et al.* 2006; Vidya *et al.* 2009). Although the mitogenome of a Columbian mammoth (*Mammuthus columbi*) has been recently sequenced, we did not include it in our alignment because the Columbian mammoth mitogenome had originated in the woolly mammoth lineage, and been transferred between the two species through hybridization (Enk *et al.* 2011).

A comparison of nucleotide differences across elephantid lineages is shown in Figure 2.2. The top panel compares elephantid mitochondrial genomes to that of forest elephant DS1534, while the bottom panel compares them to the reference savanna elephant S clade mitochondrial genome (Hauf *et al.* 1999). The mitogenomes of the *Loxodonta* individuals shown are representative of the deepest divergences present within the F and the S mtDNA clades (Debruyne 2005; Lei *et al.* 2008; Murata *et al.* 2009; Roca *et al.* 2005). It is evident that the differences between the two forest elephant F clade genomes (Figure 2.2, top panel) are greater than the differences that exist across the savanna elephant S clade mitochondrial genomes (Figure 2.2, bottom panel), consistent with an older coalescent date for F clade than for S clade mitogenomes (see below).

Phylogenetic Analyses and Molecular Dating

Phylogenetic relationships across lineages were inferred using maximum parsimony, Neighbor-Joining, minimum evolution and maximum likelihood methods. All methods inferred the same relationships among the elephantid mitogenomes. For all methods, all nodes were supported with 100% bootstrap support. A Bayesian approach inferred the same relationships

across the phylogeny (Figure 2.3) as the other methods, with a Bayesian posterior probability of 1.00 for all nodes. The results were consistent with previous molecular studies showing that woolly mammoths are closer to Asian elephants than to African elephants (Krause *et al.* 2006; Maikaew *et al.* 2007; Murata *et al.* 2009; Rogaev *et al.* 2006; Rohland *et al.* 2007; Rohland *et al.* 2010). Deep divergence was also inferred for the split between the F and S clade mtDNA lineages in African elephants (Figure 2.3).

To estimate divergence dates, we used two sets of fossil calibration dates for nodes on the elephantid phylogeny (Table 2.2). Using the "broad" fossil calibration dates (see above, and Table 2.2), the divergence between *Loxodonta* and the other elephantids was estimated as occurring 6.06 (95% CI 3.92-8.50) Mya, while the date for the *Elephas-Mammuthus* divergence was estimated as 5.42 (95% CI 3.40-7.62) Mya (Table 2.2). The divergence between the forest elephant F clade mitochondrial genomes and the S clade savanna elephant genomes was estimated as 4.86 (95% CI 2.96-6.90) Mya, almost as deep as the split between Asian elephant and woolly mammoth mitogenomes (Table 2.2). Using the "narrow" range of fossil calibration dates (see above, and Table 2.2), the divergence between *Loxodonta* and the other elephantids was estimated as 6.81 (95% CI 5.43-8.23) Mya (Table 2.2, Figure 2.3), while the *Elephas-Mammuthus* divergence was estimated as 6.01 (95% CI 4.71-7.17) Mya (Table 2.2, Figure 2.3). The divergence between African elephant F and S clades was estimated as 5.51 (95% CI 4.26-7.24) Mya, or about 92% of the divergence time estimate for the mammoth-Asian elephant split (Table 2.2, Figure 2.3).

While it had been appropriate to use the mastodon as an outgroup for inferring divergence dates among elephantid genera (Roca 2008; Rohland *et al.* 2007; Rohland *et al.* 2010), we thought that for inferring the divergence between the African elephant F and S clades, fossil calibration dates for *Loxodonta*, *Elephas* and *Mammuthus* represented more recent and thus more appropriate outgroups. Nonetheless, we considered and tested whether the addition of the mastodon mitochondrial genome (Rohland *et al.* 2007) to the elephantid mitogenomic dataset might have a great effect on divergence date estimates for the elephantids. Adding the mastodon would have the benefit of polarizing the polymorphisms present across elephantids by identifying potential ancestral character states for nucleotides. This could potentially improve the accuracy of date estimates by rooting the crown of the elephantid tree. Yet when the mastodon mitogenome and calibration dates (Table 2.2) were included in the analyses, there was no effect

on the relative estimates for the split between African elephant F and S clades as a proportion of the estimate for the split between Asian elephants and mammoth: the ratio was 0.92 when the mastodon was not included and 0.92 when the mastodon was included (Table 2.2), using the narrow fossil calibration dates. Thus rooting with the mastodon appears to have left divergence ratios unchanged.

Within-taxon Coalescent Dates

The estimate for the F clade coalescent date was 2.57 (95% CI 2.07-3.15) Mya for the narrow fossil calibration priors when the mastodon mitogenome was included in the alignment and the mastodon-elephantid calibration date was used along with fossil calibration dates for elephantid genera (Table 2.2). This compared to 2.43 (95% CI 1.68-3.38) Mya estimated using the narrow fossil date priors as the F clade coalescent date with only elephantids used in the alignment or as fossil calibrations (Table 2.2, Figure 2.3). Both estimates place the initial divergence of crown group forest elephant mitogenomes as occurring near the start of the Pleistocene.

Our mtDNA coalescent estimates (Table 2.3) were compared to the within-taxon coalescent estimates for nuclear loci determined by Rohland et al (2010). Our mtDNA coalescent estimates were based on the most diverse mtDNA clades within each lineage. Our mtDNA coalescent estimates are likely to accurately reflect basal within-lineage coalescent dates, and not represent inadequate sampling, since sampling for all of the mtDNA lineages involved has been quite extensive, involving dozens of locations and up to hundreds of individuals from across the entire range of each taxon (Fernando *et al.* 2000; Fleischer *et al.* 2001; Gilbert *et al.* 2008; Ishida *et al.* 2011b; Krause *et al.* 2006; Vidya *et al.* 2009). Likewise, nuclear coalescent estimates are likely to be accurate since they had been calculated for each taxon using 375 loci (Rohland *et al.* 2010). In order to account for different assumptions or methods used to generate the coalescent estimates, we normalized the intra-taxon coalescent dates as a percent of the divergence estimated between Asian and African elephants (Table 2.3). In populations with random progeny production and no sex difference in dispersal, mtDNA is expected to have an effective population size (and hence coalescent date) one-fourth that of nuclear loci (Hedrick 2007). Yet for each elephantid taxon, the relative coalescent for mtDNA was much higher than 0.25, and in each case greater than the coalescent for nuclear DNA (Table 2.3). The most extreme value was

estimated for the woolly mammoth, for which the coalescent, standardized to the Asian-African elephant divergence, was more than twice as old for mtDNA as for nuclear loci (Table 2.3). We also compared the mtDNA coalescent to the coalescent of nuclear loci that were two standard deviations above the mean nuclear coalescent estimate. The mtDNA:nuclear coalescent ratio was much higher than 0.25 even when the coalescent for the mtDNA was compared to the estimated coalescent for nuclear loci two standard deviations above the mean (Table 2.3).

Discussion

For a number of elephantid taxa, mtDNA phylogeographic patterns appear to be incongruent with nuclear genetic or morphological patterns (Debruyne 2005; Enk *et al.* 2011; Ishida *et al.* 2011b; Lei *et al.* 2008; Lei *et al.* 2009, 2012; Roca *et al.* 2005; Roca *et al.* 2007). One hypothesis has attributed this incongruence to sex differences in reproductive competition and dispersal (Ishida *et al.* 2011b). One prediction based on this hypothesis would be that the ratio of mtDNA coalescent dates to nuclear coalescent dates should be much greater than the 0.25 expected in populations with random progeny production and no sex difference in dispersal. In elephantids, both low female dispersal and high male variance in reproductive success would tend to increase mtDNA effective population sizes relative to those of nuclear loci (Hedrick 2007; Hoelzer 1997).

Using within-taxon coalescent estimates for nuclear markers estimated by Rohland et al. (2010), we were able to determine that mtDNA coalescent dates were indeed older than those of nuclear loci (Table 2.3). The coalescent data estimates for nuclear data are likely to accurately reflect values for neutral loci in the nuclear genome since: (1) sequences had been randomly generated, so were likely to be non-coding loci not under selection; and (2) repetitive elements had been screened out, so that sequences across taxa were likely to represent orthologous and not paralogous loci (Rohland *et al.* 2010). Elephantid mtDNA coalescent dates were very high relative to nuclear coalescent dates even when the mtDNA coalescent was compared to the estimated coalescent for nuclear loci that are two standard deviations above the mean coalescent (Table 2.3). Thus our study provides support for the hypothesis that among elephantids coalescent dates for mtDNA relative to nuclear DNA would be much older than the 0.25

expected in populations with random progeny production and no sex difference in dispersal. The relatively ancient mtDNA coalescent dates are consistent with lower female than male dispersal and reproductive competition in elephantids, which would allow mtDNA genetic patterns to locally persist even as male dispersal would disrupt any geographic subdivisions among nuclear alleles (Ishida *et al.* 2011b; Petit & Excoffier 2009; Wright 1943).

The effects of sex differences in reproductive success on effective population size (Ne) have been previously quantified (Hedrick 2007; Hoelzer 1997). As male effective population sizes are reduced, Ne drops for nuclear loci (Hedrick 2007; Hoelzer 1997). The effective population size becomes equal for autosomal, X-linked and mtDNA loci when the male effective population size is reduced to one-seventh that for females (Nem = Nef/7), i.e., when the effective population size for males comprises one-eighth that of the total effective population size of males and females combined (Hedrick 2007; Hoelzer 1997). Such a ratio could be consistent with the lack of reproductive success documented for a large proportion of males in field studies of savanna elephants (Hollister-Smith *et al.* 2007; Poole *et al.* 2011). For example, in a 22-year study at Amboseli National Park, of 89 male elephants genotyped, 53 (60%) were not found to have sired a calf; while 30% of the 119 calves examined had been fathered by just three males (Hollister-Smith *et al.* 2007). Male savanna elephants represent the extreme among mammals in the extent to which high mating and paternity success occur late in life (Hollister-Smith *et al.* 2007). Age-specific paternity peaks at 45-53 years of age (Hollister-Smith *et al.* 2007). Since fewer than 10% of males are estimated to survive to age 50 (Poole *et al.* 2011), a high proportion of males are not reproductively successful, and this high variance in reproductive success would reduce the effective population size of nuclear markers (Hedrick 2007; Hoelzer 1997).

Likewise, the lower degree of dispersal for female than male elephants would be expected to increase the effective population size and coalescent of mtDNA relative to those of nuclear loci (Hoelzer 1997; Ishida *et al.* 2011b; Wright 1943). The interspecies transfer of F clade mtDNA from forest to savanna elephant populations (Debruyne 2005; Ishida *et al.* 2011b; Roca *et al.* 2005) was also likely enhanced by low rates of female dispersal (Petit & Excoffier 2009), although we could disregard the effects of hybridization by considering only S clade coalescent dates among savanna elephants. When dispersal rates for males are high while females exhibit high matrilocality, the sex-biased dispersal can increase the effective population

size of mtDNA relative to nuclear DNA even to the point of causing the coalescent of mtDNA to exceed that of nuclear loci (Hoelzer 1997).

For both nuclear and mtDNA loci, coalescent estimates are low for the savanna elephant (low diversity and effective population sizes relative to other taxa); high for the forest elephant (highest diversity and effective population size); and intermediate for the Asian elephant (intermediate diversity and effective population size) (Table 2.3) (Rohland *et al.* 2010). Thus the rank order of genetic diversity for these three lineages was the same for mtDNA and nuclear markers (Table 2.3). The savanna elephant has the lowest nuclear coalescent among the elephantids (Rohland *et al.* 2010), and the mtDNA coalescent date for the savanna elephant S clade was also low relative to other taxa (Table 2.3). A number of authorities have suggested that the geographic range and numbers of savanna elephants expanded towards the end of the Pleistocene following the extinction of *Elephas*, which had been the predominant elephantid on the African savannas until that time (Kingdon 1979; Maglio 1973; Sanders *et al.* 2010). The relatively low coalescent values for both nuclear and S clade mtDNA in the savanna elephant (relative to other taxa) could reflect this proposed Late Pleistocene founder effect for the savanna elephant. Simulations and empirical data have shown that the interspecies transfer of alleles to a species is expected to occur as its dispersing sex colonizes a region occupied by the other species, with the other species transmitting markers carried by the non-dispersing sex to the colonizing species (Currat *et al.* 2008; Petit & Excoffier 2009). The transfer of F clade mtDNA from forest to savanna elephants would be consistent with this expectation, if male savanna elephants had expanded their range into formerly forest elephant habitats (Roca *et al.* 2005). The F clade mtDNA sequences shared by forest and savanna species are quite similar (Eggert *et al.* 2002; Roca *et al.* 2005), indicating that the interspecies transfer would be consistent with a Late Pleistocene event (Murata *et al.* 2009).

For the woolly mammoth, the level of nuclear diversity was low and comparable to that of savanna elephants. Yet the mtDNA coalescent date was much older for mammoths than for S clade savanna elephants (Table 2.3). The low nuclear genetic diversity present in the woolly mammoth could reflect high (relative to other species) levels of male reproductive competition (Rohland *et al.* 2010). Male-male competition often also leads to the evolution of sexual size dimorphism (Weckerly 1998), which is also observed in mammoths, with males about one meter taller than females (Haynes 1991). The woolly mammoth mitochondrial coalescent seems

especially ancient relative to the nuclear coalescent estimate (Table 2.3), and the mammoth was the only species for which the mtDNA coalescent estimate (Table 2.2, Figure 2.3) was more than twice as old as those previously reported for nuclear loci (Rohland *et al.* 2010). This pattern may reflect a more ancient widespread distribution for this species, its survival in multiple glacial refugia, or the effects of the distribution of woolly mammoths across two continents that were only intermittently connected during Pleistocene glacial cycles (Barnes *et al.* 2007; Debruyne *et al.* 2008; Gilbert *et al.* 2008; Miller *et al.* 2008). These factors would be more likely to affect mtDNA than nuclear genetic patterns among elephantids since (1) male-male competition would lower effective population size for nuclear loci, reducing overall nuclear genetic diversity but not affecting mtDNA diversity, and (2) male-mediated gene flow would tend to diminish nuclear genetic differences across populations, while not affecting mtDNA differences across populations.

As part of this study, we conducted the first full sequencing, to our knowledge, of mitogenomes from forest elephants. These were, exclusive of the control region, 15,418 bp in length for DS1534, and 15,420 bp for SL0001. The size difference was due to an insertion of 2 bp in the 12S rRNA of SL0001, not found in any other proboscidean mitogenome. The previously published savanna elephant S clade mitochondrial genome reference sequence was 15,418 bp in length (exclusive of control region) (Hauf *et al.* 1999). Since some savanna elephants carry F clade mitochondrial genomes (Debruyne 2005; Roca *et al.* 2005), savanna elephant nuclear genes involved in mitochondrial function would appear to be capable of interacting with mtDNA-coded proteins whether these derive from F clade or S clade mitochondrial genomes. Thus amino acid differences between forest F and previously sequenced savanna S clade mtDNAs (Figure 2.1) would not be expected to have led to greatly altered protein function (McKenzie *et al.* 2003), although an alternative is possible. Genetic markers carried only by the non-dispersing sex are expected to be less subject to the effects of selection (Petit & Excoffier 2009). This is somewhat paradoxical, given that the higher effective population sizes among such markers (Hoelzer 1997; Wright 1943) would make fewer mutations effectively neutral (Ohta 1973). However, limited gene flow can preclude selective sweeps within species at loci carried only by the non-dispersing sex (Petit & Excoffier 2009). This may suggest that F clade mitogenomes in savanna elephants could persist even should they lead to somewhat lower levels of fitness than S clade mitogenomes.

Our results also confirmed previous reports that mammoths are closer to living Asian elephants than to African elephants (Krause *et al.* 2006; Maikaew *et al.* 2007; Rogaev *et al.* 2006; Rohland *et al.* 2007; Rohland *et al.* 2010). The split between *Elephas* and *Mammuthus* occurred 6.0 (95% CI 4.71-7.17) Mya, while *Loxodonta* F and S clades diverged 5.5 (95% CI 4.26-7.24) Mya (Figure 2). Our estimate for the divergence of F clade and S clade African elephant mitochondrial lineages is comparable to or somewhat older than previous estimates for the divergence between forest and savanna elephants (Murata *et al.* 2009; Roca *et al.* 2005; Roca *et al.* 2001; Rohland *et al.* 2007), and suggests that the four major elephantid mtDNA lineages were established around the end of the Miocene Epoch. In Africa during this period there were cooler and more arid climates, changes in vegetation marked by a reduction in forests, and shifts in habitat zones (Bonnefille 2010; Cerling *et al.* 1998; Cerling *et al.* 1997; Griffin 2002; Maley 2001; Otero *et al.* 2011; Sepulchre *et al.* 2006). These changes may have driven the evolution not only of elephantid lineages but also of other modern mammalian lineages, including artiodactyls, perissodactyls, and primates including hominids (Cerling *et al.* 1998; Folinsbee & Brooks 2007; Roca 2008; Werdelin & Sanders 2010).

The basal divergence represented by the two sequenced forest elephant mtDNA genomes is estimated to have occurred 2.43 (95% CI 1.68-3.38) Mya, around the beginning of the Pleistocene Epoch 2.58 Mya (as recently defined to include the Gelasian Stage) (Gibbard & Head 2010). Divergence between humans and chimpanzees is comparable to the divergence across elephantid taxa, yet the coalescent date for all living human mitogenomes is under 200,000 years (Ingman *et al.* 2000). Previous studies have shown that mtDNA differences persist in elephants within and between geographic regions even in the face of changes in habitats or in nuclear genotypes (Ishida *et al.* 2011b; Roca *et al.* 2005; Roca *et al.* 2007; Roca & O'Brien 2005). A close association of elephant mtDNA lineages with geographic region would be expected since females do not typically migrate away from natal social groups. Since lack of female dispersal increases the effective population size of mtDNA (Table 2), mtDNA lineages would be removed by genetic drift much less rapidly than would nuclear diversity, so that geographic partitions caused by ancient climate and habitat changes would be detectable in the mtDNA phylogeographic patterns long after the nuclear phylogeographic partitions had been erased by male-mediate gene flow (Hoelzer 1997; Petit & Excoffier 2009; Wright 1943).

The deepest subdivision within the forest elephant F clade mtDNA could reflect the initial fragmentation of African tropical forest habitats during glacial cycles that began 2.5 Mya (Figure 2.3) (Maley 2001). With the caveat that locations and definitions of Pleistocene refugia for tropical forests remain controversial (Lowe *et al.* 2010), African tropical forest habitats are believed to have repeatedly fragmented during glacial cycles into a set of discontinuous refugia (DeMenocal 2004; Maley 2001; Mayr & Ohara 1986). Forest elephant populations within refugia may have been separated by intervening savanna habitats occupied by larger elephantid species (Kingdon 1979; Maglio 1973; Sanders *et al.* 2010). The two forest elephant mitogenomes sequenced were from an individual from the West African Guinean forest block and an individual from the Central African Congolian forest block, although it is not clear whether this deepest F clade subdivision was consequent to an east-west separation of the tropical forest, as currently occurs (Salzmann & Hoelzmann 2005; White 1983). Elephants carrying mtDNA haplotypes from each of the two deepest subdivisions of the F clade are currently present in populations in both West and Central Africa, an indication that some migration of elephant females across the two forest blocks occurred subsequent to the basal F-clade split (Debruyne *et al.* 2003b). Also, given that gene flow in elephants is male-mediated and that the long-distance dispersal of forest elephant nuclear alleles is well documented (Roca *et al.* 2005; Roca *et al.* 2001), the partitioning of mtDNA would not necessarily imply the establishment or persistence of nuclear genetic differentiation between the two forest blocks (Ishida *et al.* 2011b; Roca *et al.* 2007). Nonetheless, since the ancient coalescent for mtDNA was likely consequent to relatively low levels of competition and limited dispersal among matrilocal females, elephant mtDNA would be likely to persistently record a phylogeographic signature of ancient climate transitions and habitat changes, such as occurred in Africa at the start of the Pleistocene.

Figures and Tables

Figure 2.1. Amino acid alignments for mitochondrial proteins across elephantids.
The translated codons of 13 protein coding genes in elephantid mitogenomes were aligned for two forest elephants (DS1534 and SL0001) and seven other elephantids (GenBank numbers: NC_000934, AB443879, DQ316069, AJ428946, NC_005129, EU153453, and NC_007596) using Geneious v5.4 (Biomatters Limited). Conserved amino acids are shown as dots below a consensus sequence. Colors indicate the hydrophobicity of amino acid residues, with red being highly hydrophobic and blue highly hydrophilic. With few exceptions the differences between the forest elephants and the other elephantids fell into the following categories: (1) amino acid differences that did not involve change from a highly hydrophobic to a highly hydrophilic residue sequence, or vice versa; (2) amino acid substitutions that were also present in at least one other elephantid species; or (3) alternative stop codon present in at least one other elephantid species. The following comprise the only exceptions: For forest elephant DS1534, ATP6 appears to rely on an alternative start codon from the other elephantids, ND4L has a different stop codon from the other elephantids, and ND4 has Gln instead of Leu in the second amino acid position. For forest elephant SL0001, ND3 relies on an alternative start codon from the other elephantids. Both forest elephants have Tyr instead of His at position 312 of CYTB. Abbreviations: LCY: *Loxodonta cyclotis*, African forest elephant; LAF: *Loxodonta africana*, African savanna elephant; EMA: *Elephas maximus*, Asian elephant; MPR: *Mammuthus primigenius*, woolly mammoth.

Figure 2.1. Continued. NADH dehydrogenase subunit 1 (ND1)

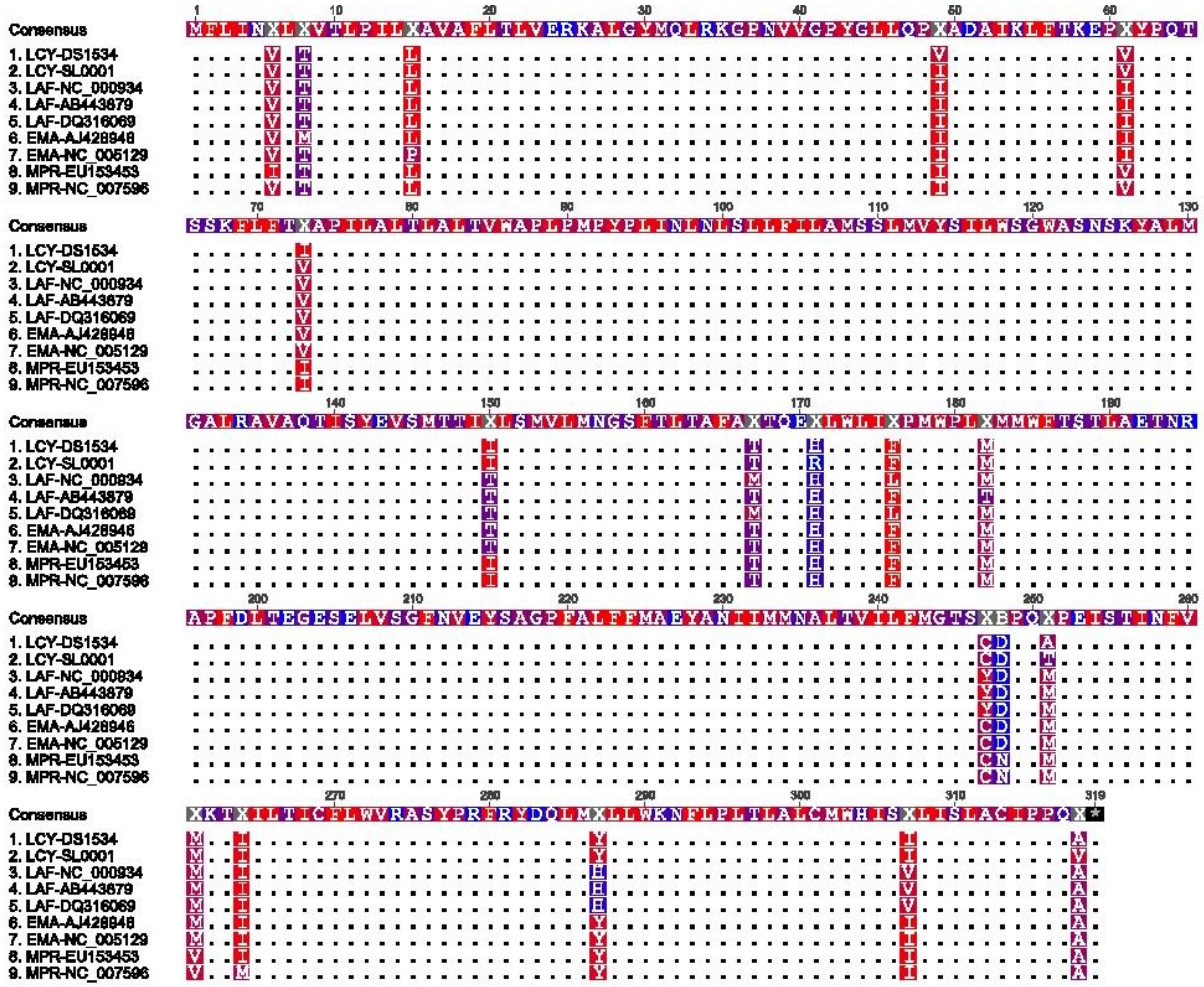


Figure 2.1. Continued. NADH dehydrogenase subunit 2 (ND2)

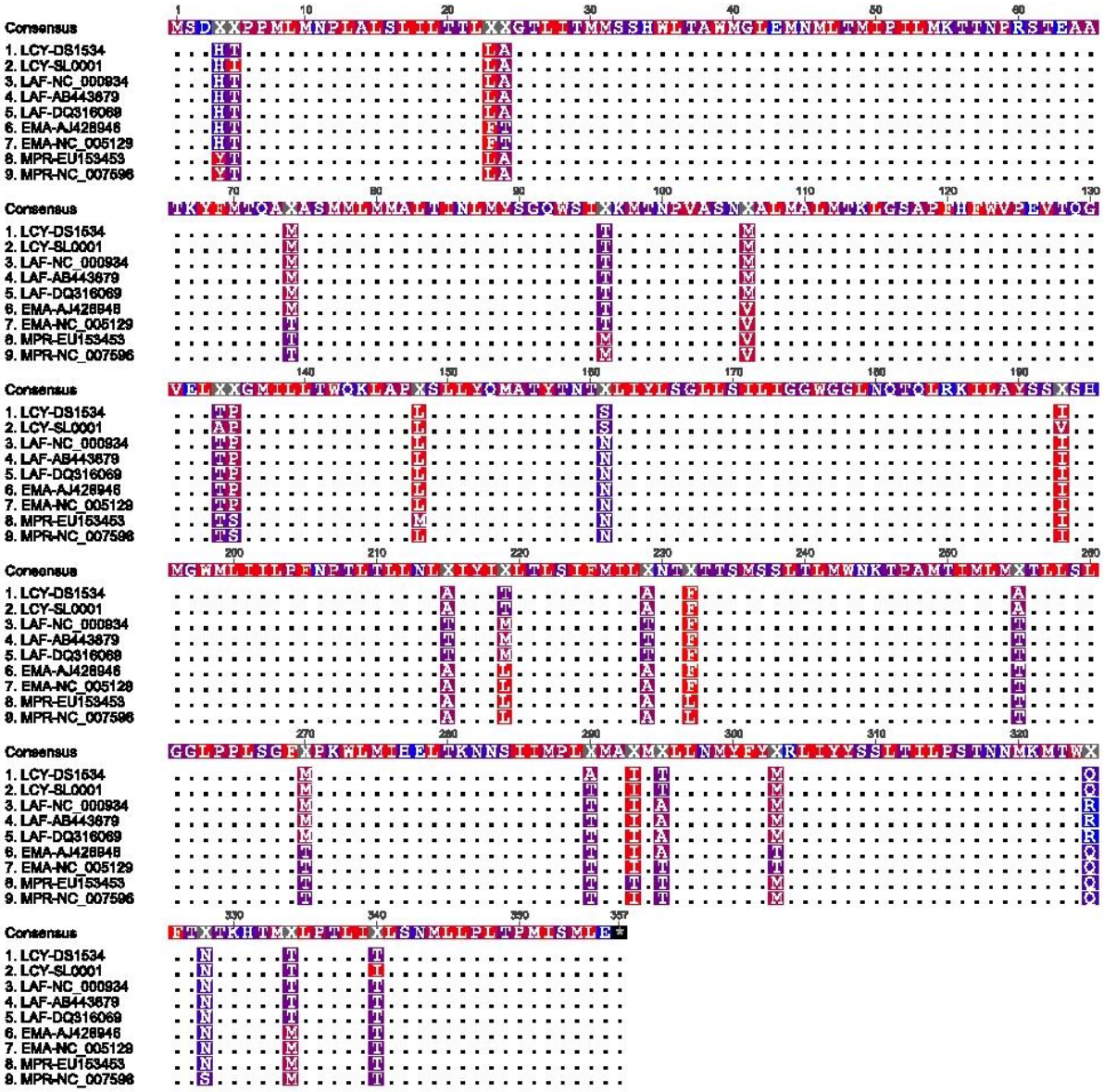


Figure 2.1. Continued. Cytochrome c oxidase subunit I (COX1)

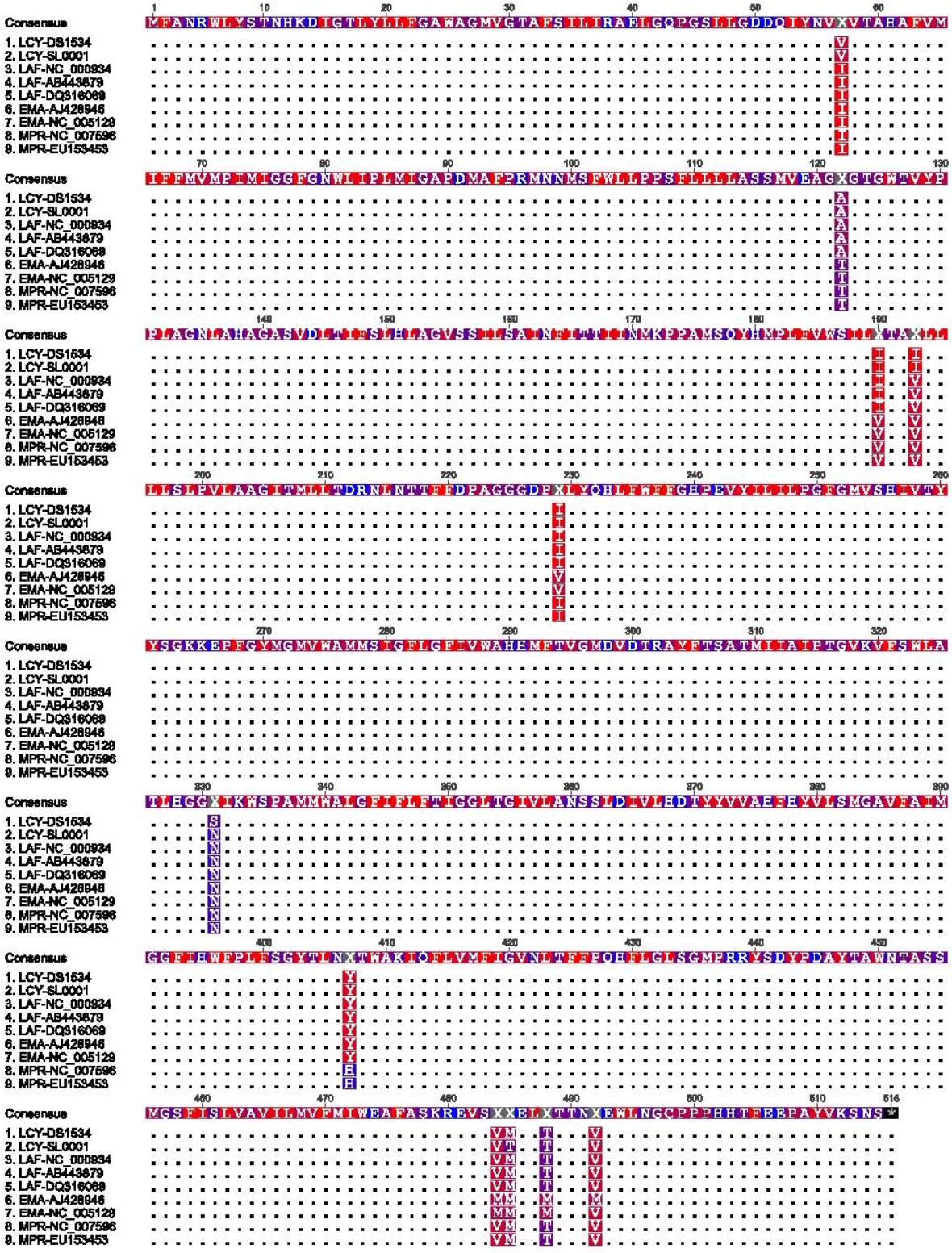


Figure 2.1. Continued. Cytochrome c oxidase subunit II (COX2)

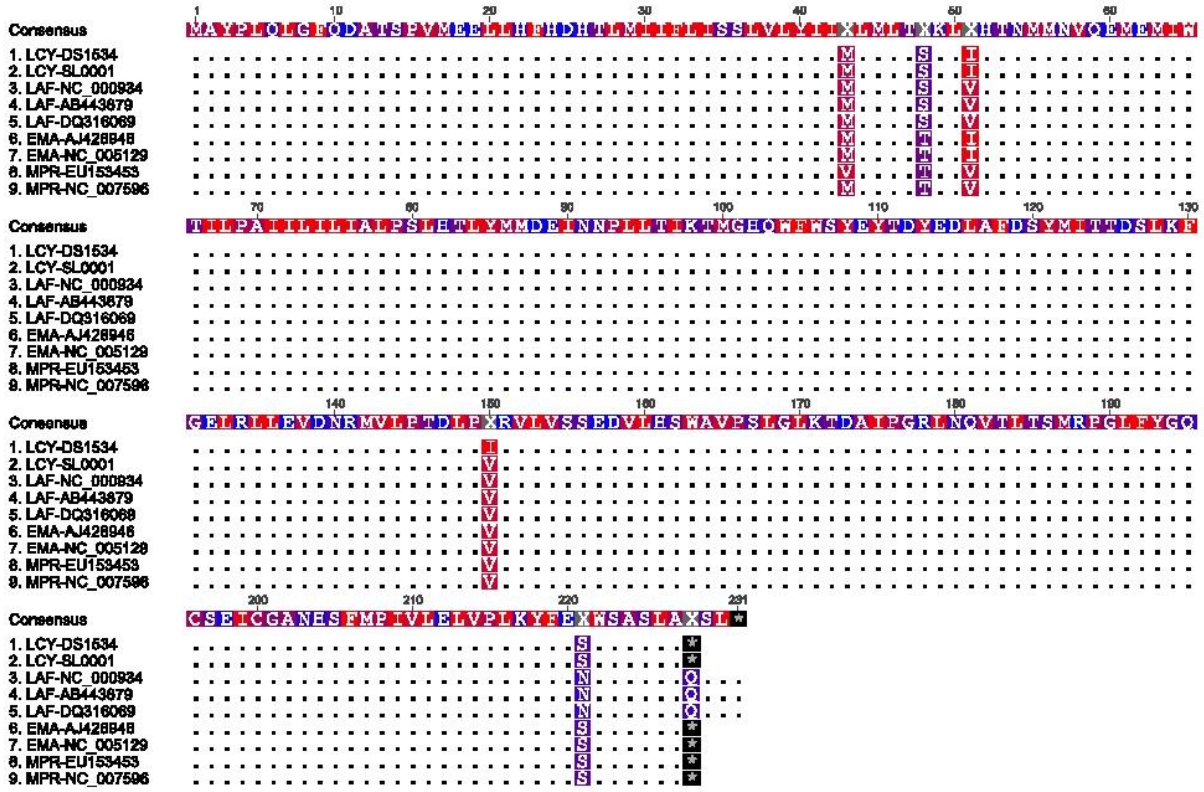


Figure 2.1. Continued. ATP synthase F0 subunit 8 (ATP8)

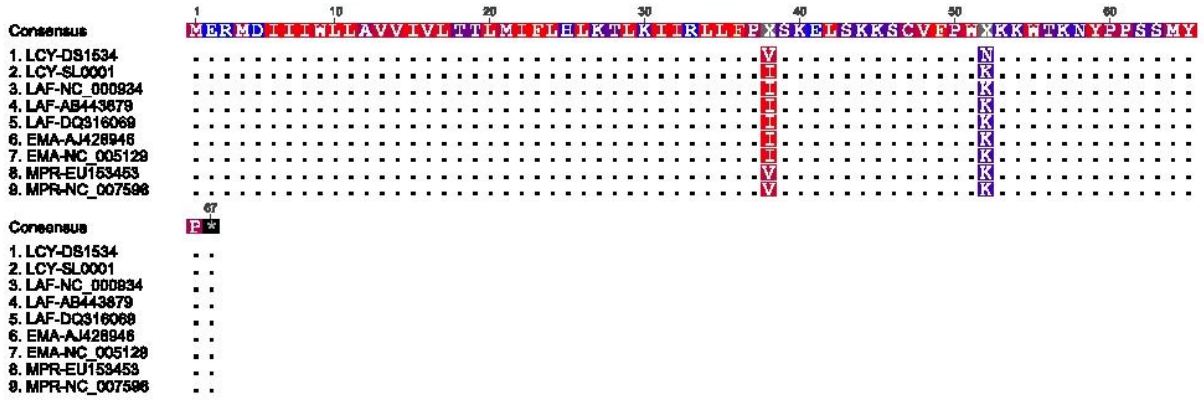


Figure 2.1. Continued. ATP synthase F0 subunit 6 (ATP6)

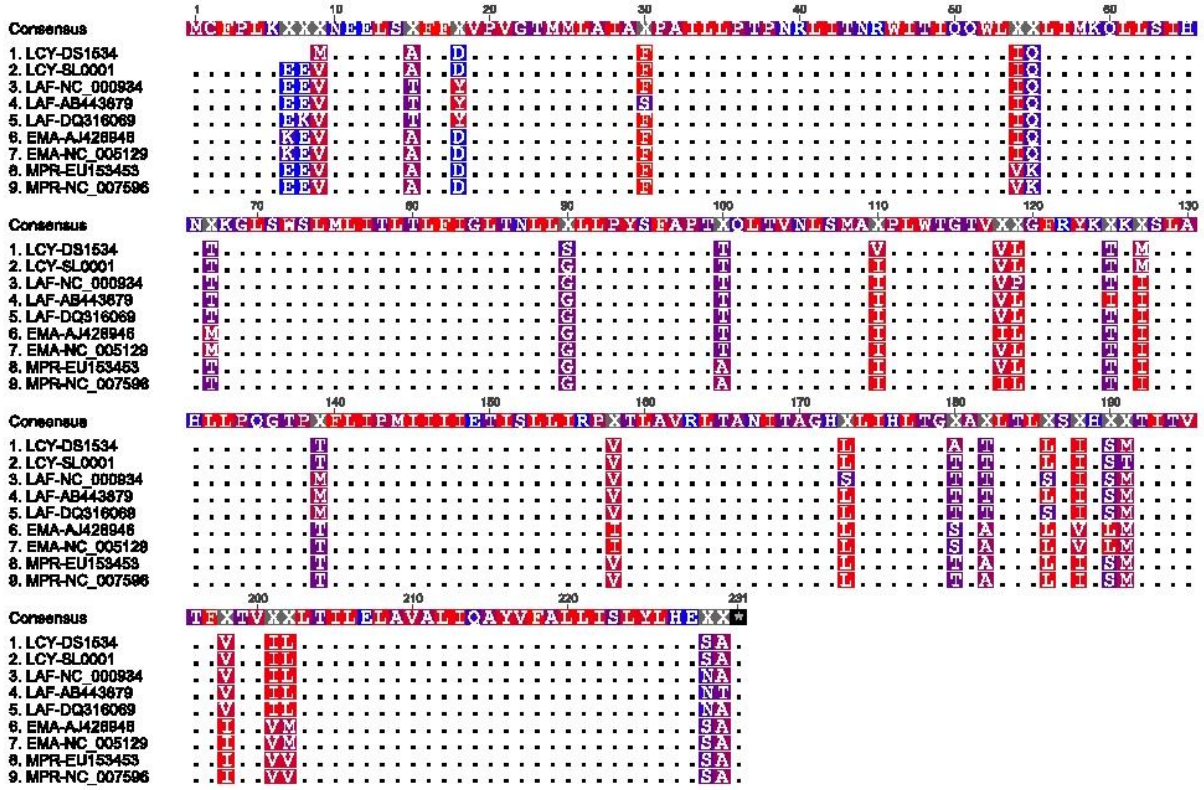


Figure 2.1. Continued. Cytochrome c oxidase subunit III (COX3)

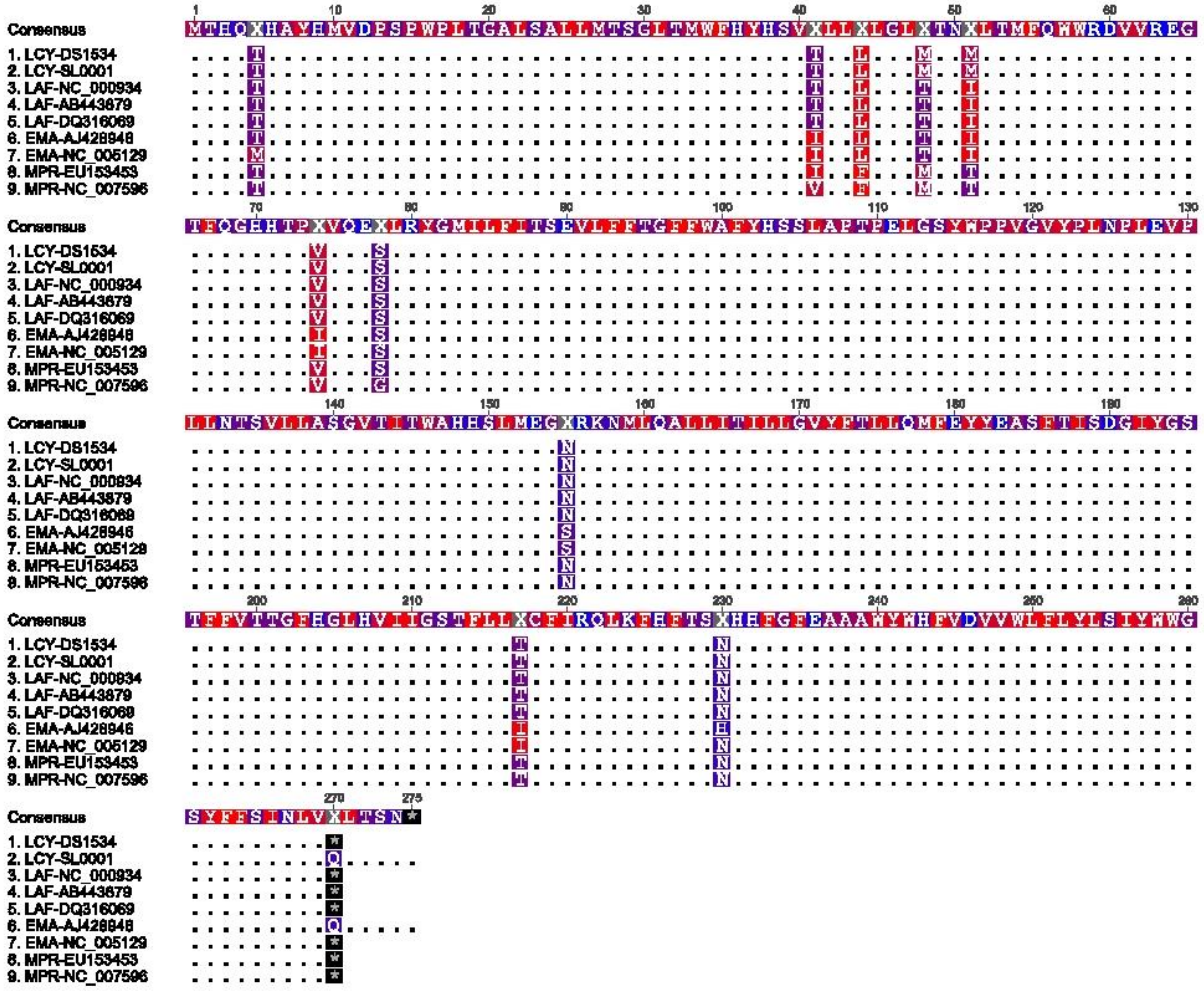


Figure 2.1. Continued. NADH dehydrogenase subunit 3 (ND3)

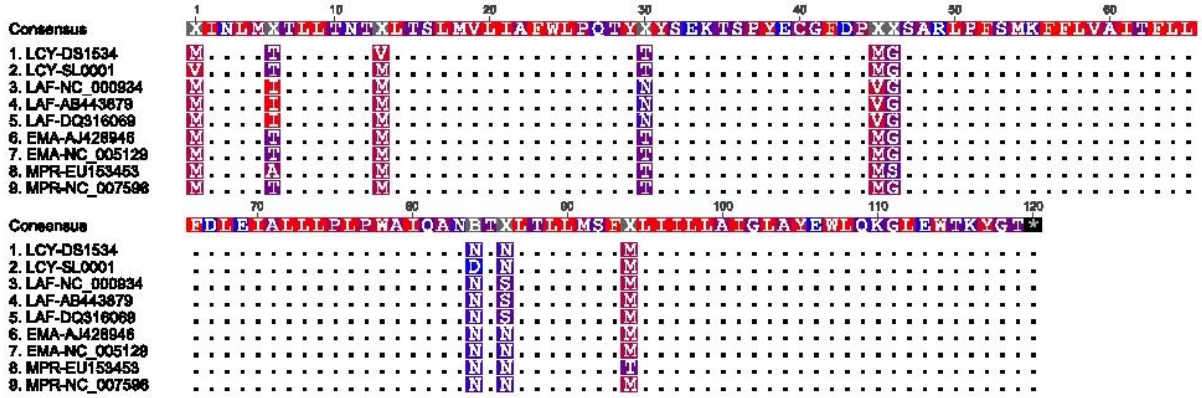


Figure 2.1. Continued. NADH dehydrogenase subunit 4L (ND4L)

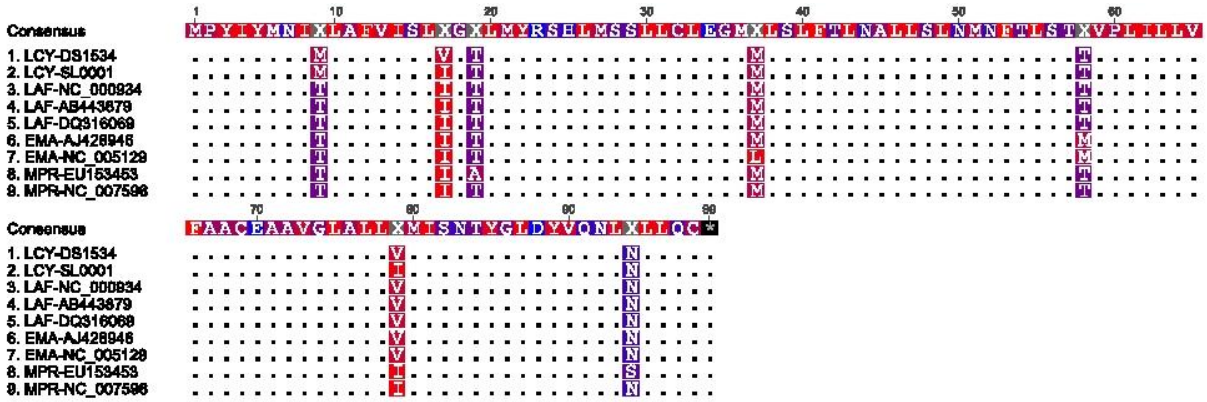


Figure 2.1. Continued. NADH dehydrogenase subunit 4 (ND4)

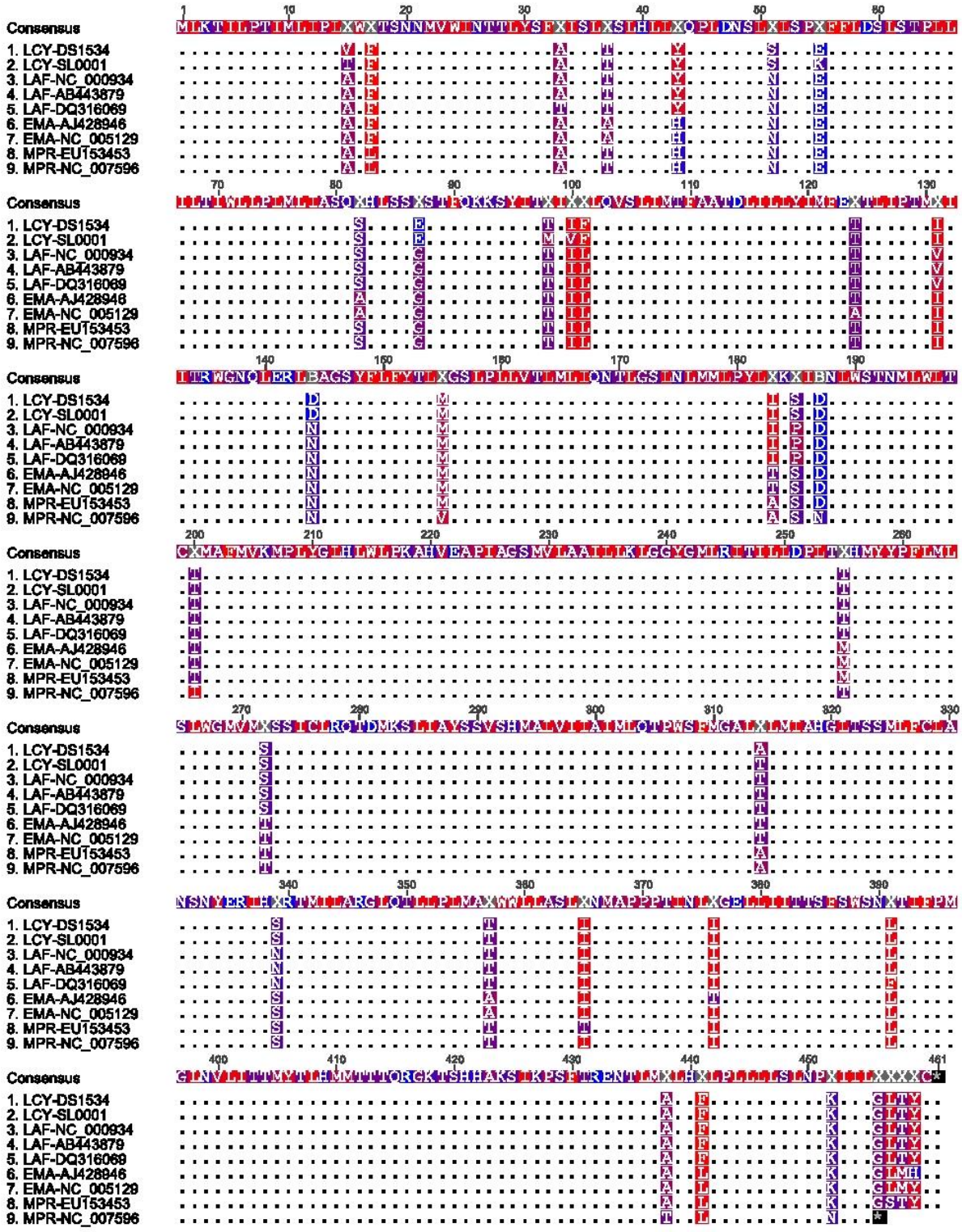


Figure 2.1. Continued. NADH dehydrogenase subunit 5 (ND5)

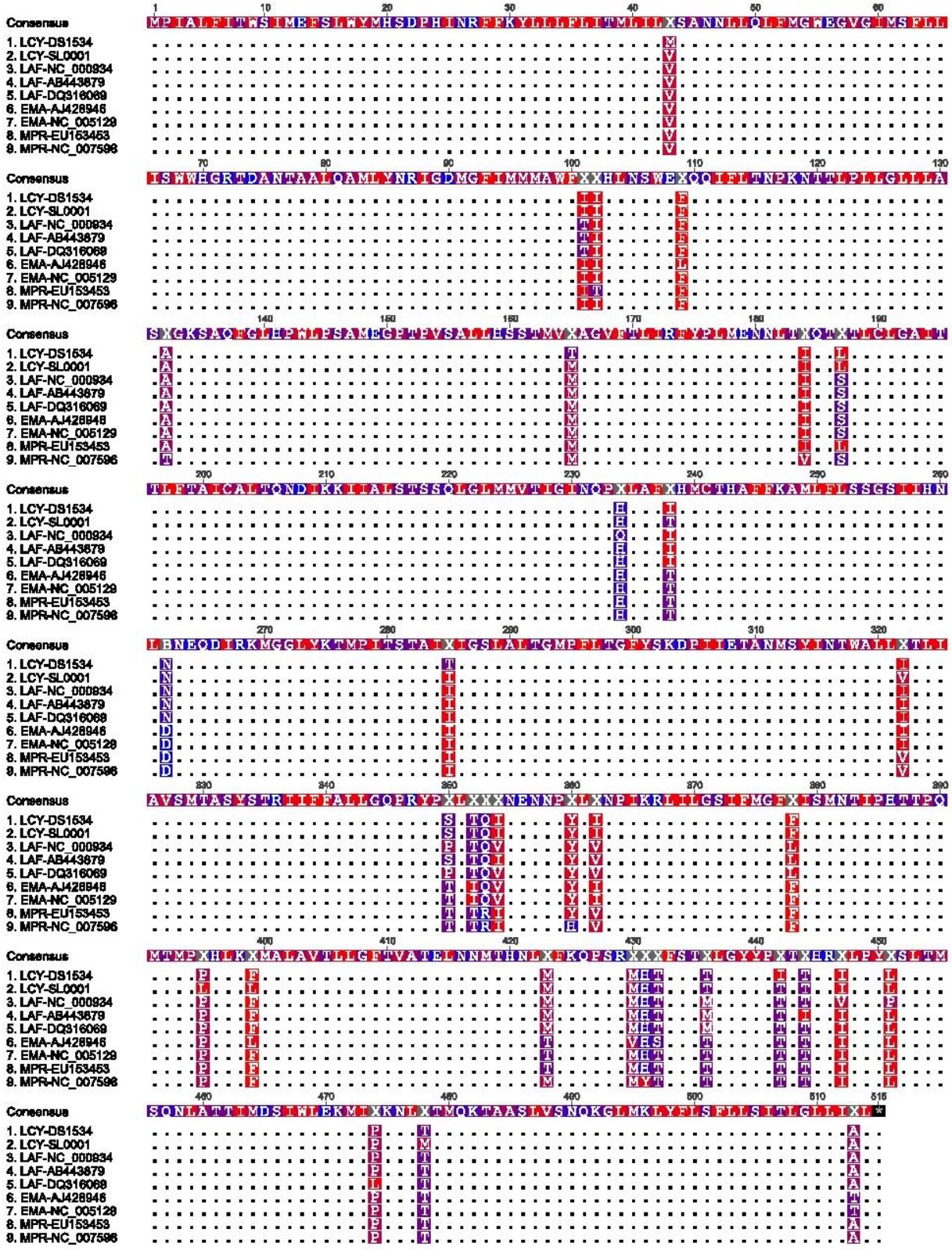


Figure 2.1. Continued. NADH dehydrogenase subunit 6 (ND6)

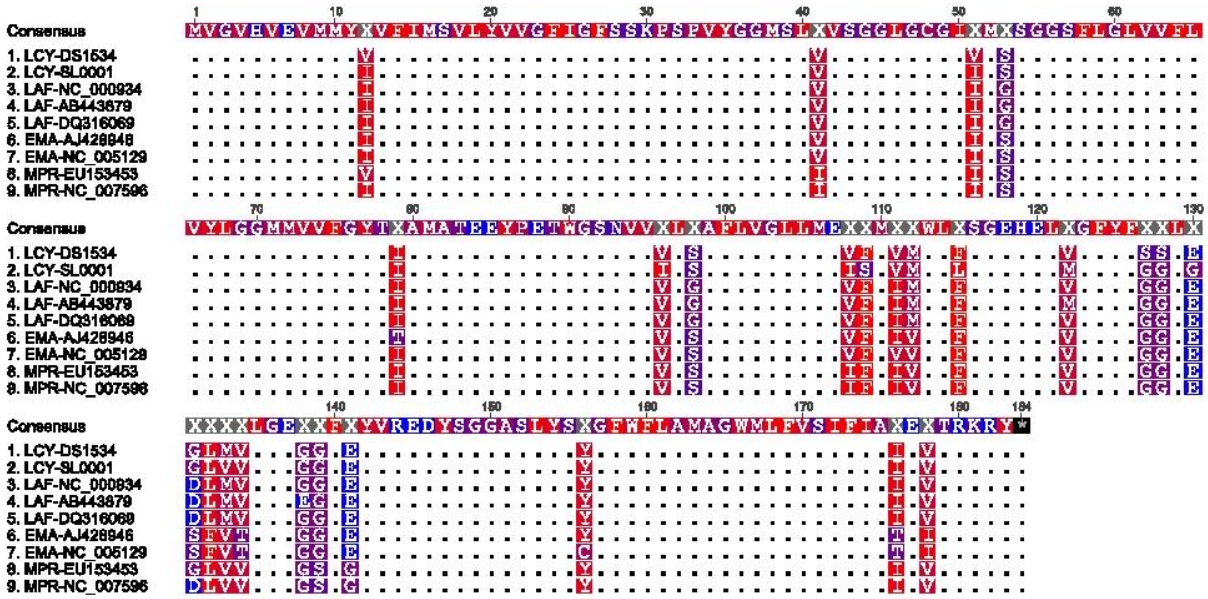


Figure 2.1. Continued. Cytochrome b (CYTB)

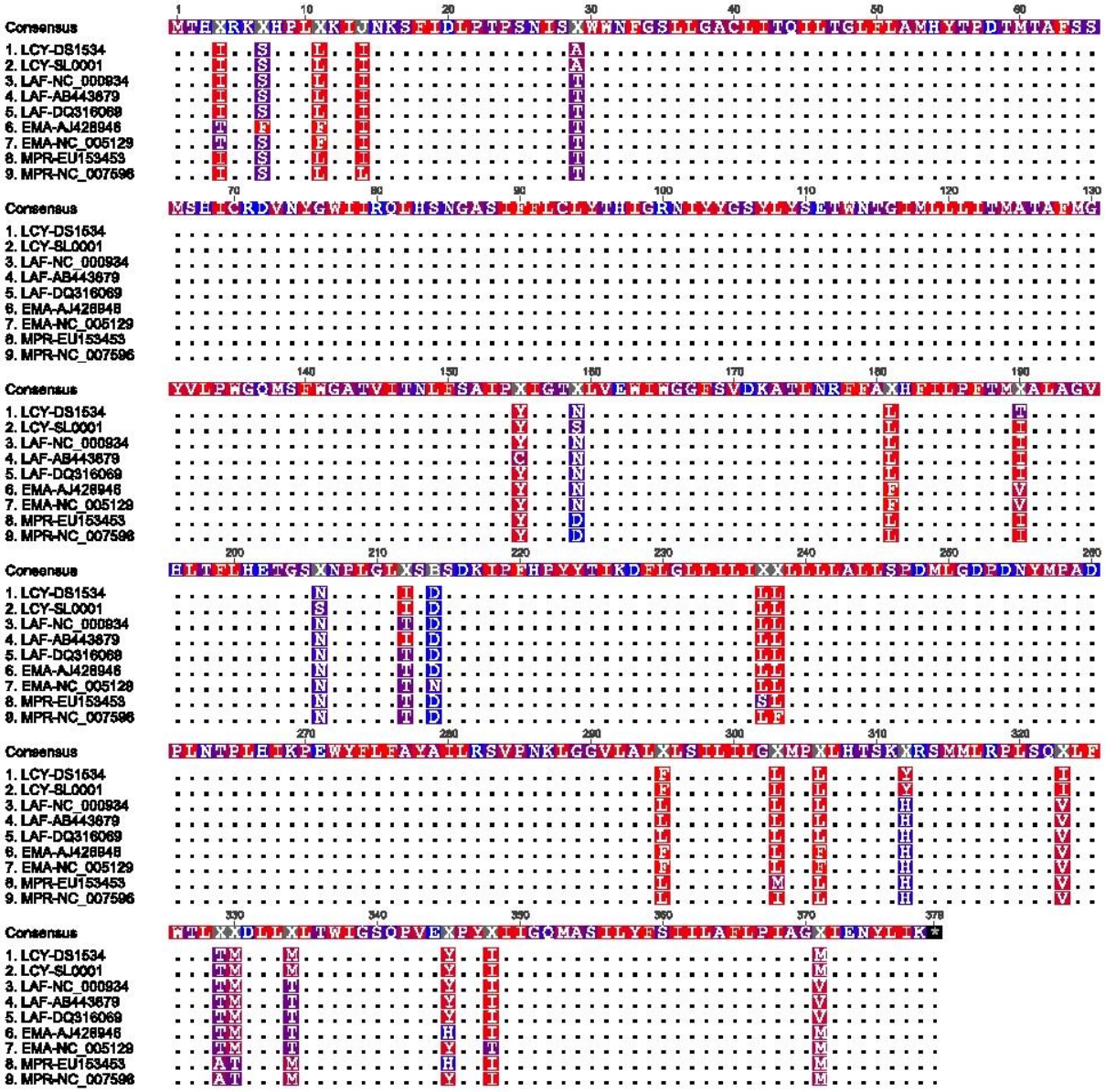


Figure 2.2. Sequence differences found among the mitochondrial genomes of elephantids with respect to African forest (LCY) or savanna (LAF) elephants.

Nucleotide comparisons between a reference genome and those of other elephantids are shown by horizontal bars in the top and bottom panels, with nucleotide matches represented by the white background, substitutions indicated by black vertical lines, and indels by horizontal lines or dots. *Top panel:* comparison of elephantid mtDNA genomic sequences to reference forest elephant DS1534. Deep within-species differences are evident in the comparison to SL0001. *Center panel:* nucleotide positions within the mtDNA genome and locations of mtDNA regions. The D-loop was excluded due to saturation of sites. *Lower panel:* comparison of elephantid mtDNA genomic sequences to reference savanna elephant NC_000934 (Hauf *et al.* 1999). Note that differences present across the savanna elephant S clade (DQ316069 and AB443879) are not as great as the differences across forest elephants shown in the top panel, reflecting the more recent coalescent date for savanna elephant S clade genomes. Panels were drawn using Geneious v5.4 (Drummond *et al.* 2010). Abbreviations: LCY: *Loxodonta cyclotis*, African forest elephant; LAF: *Loxodonta africana*, African savanna elephant; MPR: *Mammuthus primigenius*, woolly mammoth; EMA: *Elephas maximus*, Asian elephant.

Figure 2.2. Continued.

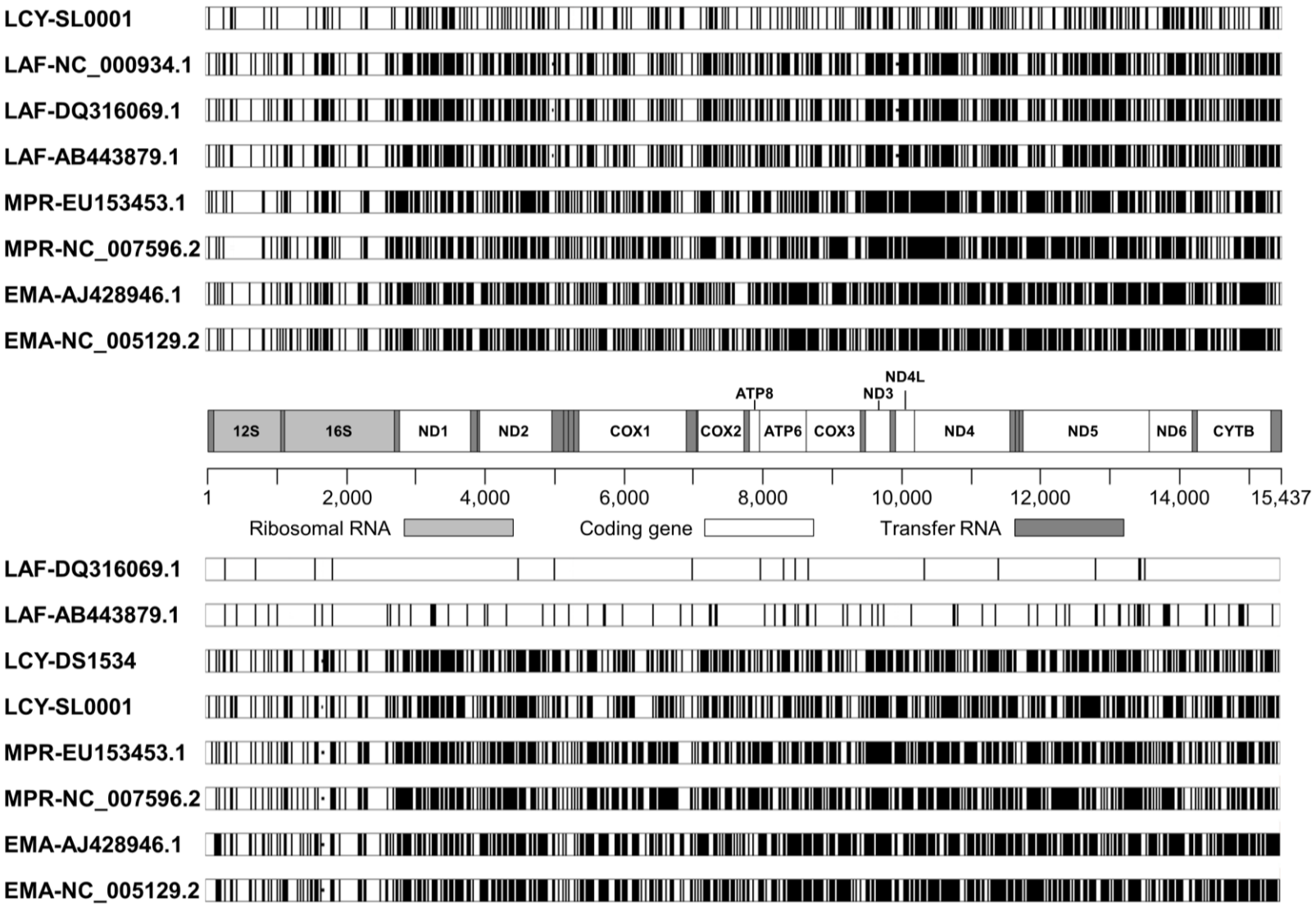


Figure 2.3. Maximum clade probability tree inferred from the alignment of elephantid mitochondrial genome sequences.

The software *BEAST* (Drummond & Rambaut 2007) was used to generate the tree, which is displayed as a chronogram. Shaded bars indicate the 95% highest posterior density as determined using two sets of calibrations. The top bar used the broadest range of calibrations based on fossil estimates (Rohland *et al.* 2010; Sanders *et al.* 2010); the lower bar relied on "narrow" set of calibration dates that excluded fossils identified as being of questionable assignment to genera by Sanders et al. (2010), and assumed the monophyly of elephantid genera (Maglio 1973; Rohland *et al.* 2010; Shoshani & Tassy 1996). The tree shown was based on the narrow fossil calibrations. The divergence date between forest elephant (F clade) and savanna elephant (S clade) mtDNA genomes was comparable to the divergence between Asian elephant and mammoth mtDNA genomes. For all four lineages, the two or three genomes shown encompass the most basal within-lineage divergences; these indicate that forest elephants have the deepest within-lineage crown-group coalescent date, while S clade savanna elephants have the most recent mtDNA coalescent date. The posterior probability was 1.00 for all nodes (leftmost number). The same relationships across mitogenomes were inferred in PAUP* (Swofford 2002) using maximum parsimony (MP), Neighbor Joining (NJ), minimum evolution (ME) and maximum likelihood (ML) methods (not shown). Bootstrap support was 100% for each method at each node (shown respectively from left to right, from the second number at each node, for MP, NJ, ME and ML). Tree scores were: MP, 1611 variable sites, 1106 parsimony informative, 1 tree, length=1867, CI=0.876, RC=0.764; ME, 1 tree, score=0.12588; ML, 1 tree, -ln likelihood=31041.20336. A time scale and geological epochs are shown below the phylogeny. Africa in the late Miocene experienced drier climates and the retreat of forests (Bonnefille 2010), and shifts in habitat zones during the Messinian (Griffin 2002). Shown also is a Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records, which measure global ice volume and deep ocean temperature (lower values reflect lower temperatures), and which illustrate the initiation of glacial cycles at the start of the Pleistocene (Lisiecki & Raymo 2005).

Figure 2.3. Continued.

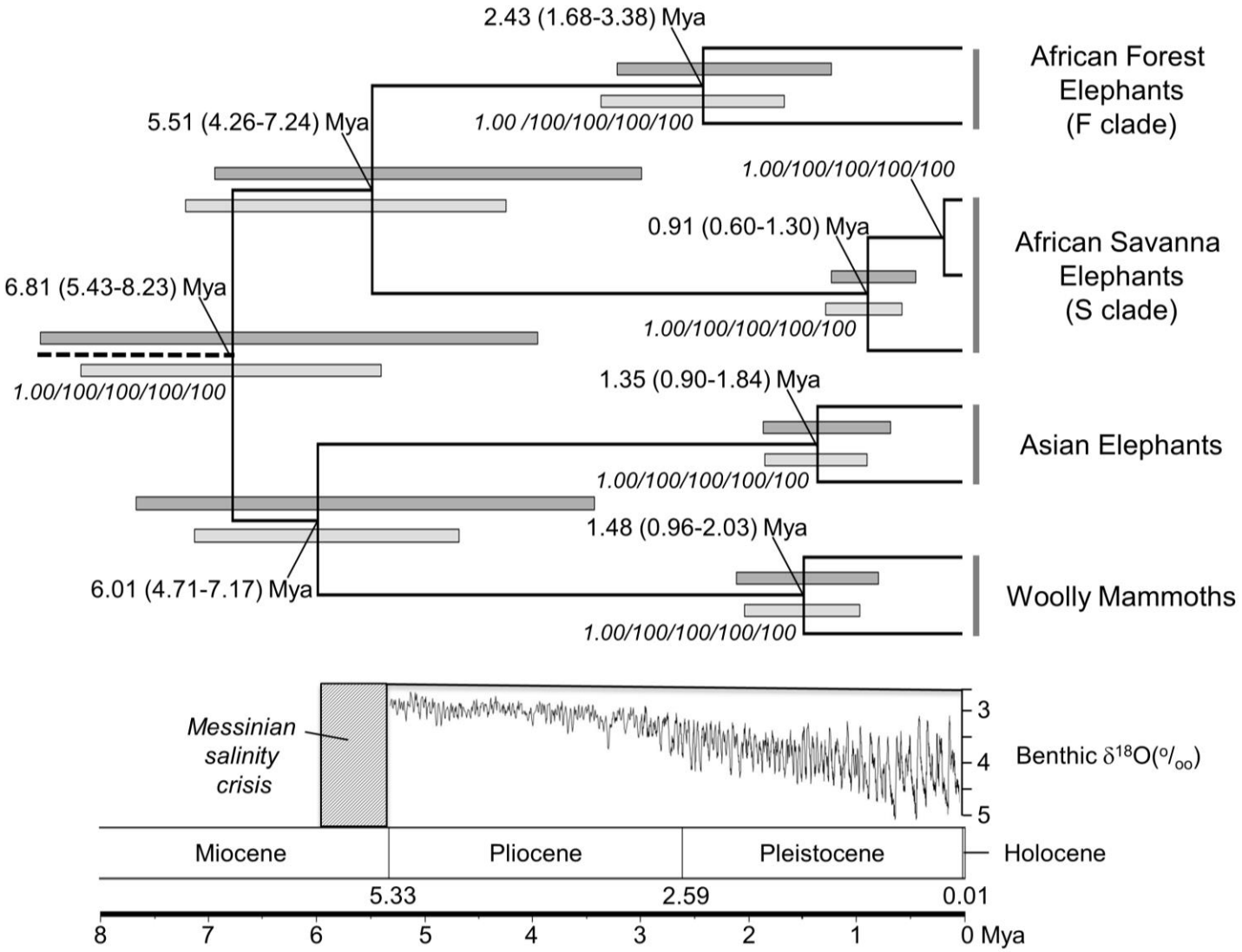


Table 2.1. Oligonucleotide primers used for PCR and sequencing of elephant mitochondrial DNA (mtDNA).

Primer name	Primer function	Primer sequence	mtDNA regions (full or partial):	Size (bp)
Set 1				
mtAll-F1	PCR, Sequencing	CCCCCACGGGAWACAGCAG	12S ribosomal RNA tRNA-Val 16S ribosomal RNA	2282
mtAll-F12	Sequencing	AAGTGAGCTTAATCATAACCCATGA		
mtAll-F13	Sequencing	AATAGAGATAGTACCGCAAGGGAAT		
mtAll-F14	Sequencing	CACTGCCTGCCCAGTGACCAAG		
mtAll-R1	PCR, Sequencing	GTCTGAACTCAGATCACGTAGGACT		
mtAll-R11	Sequencing	CGGTCTGAACTCAGATCACGTAG		
mtAll-R12	Sequencing	GAGACAGTTAAACCCTCGTGTG		
mtAll-R13	Sequencing	TGAATTTACTCGCTAATCAAGGTTGT		
Set 2				
mtAll-F2	PCR, Sequencing	GACCTCGATGTTGGATCAGG	16S ribosomal RNA tRNA-Leu NADH dehydrogenase subunit 1 (*ND1*) tRNA-Ile tRNA-Gln tRNA-Met NADH dehydrogenase subunit 2 (*ND2*) tRNA-Trp tRNA-Ala tRNA-Asn tRNA-Cys	2821
mtAll-F21	Sequencing	CCCCAATTCTAGCCCTAACC		
mtAll-F22	Sequencing	GAAACTAACCGAGCCCCATT		
mtAll-F23	Sequencing	CGAACCTAAACTCGAGAATTCAA		
mtAll-F24b	Sequencing	CACCTACACCAACACCAACCT		
mtAll-R2	PCR, Sequencing	TGCAAATTCRAAGAAGCAG		
mtAll-R21	Sequencing	AAATTATTGGGGTTAGGGGAAG		
mtAll-R22	Sequencing	TGTTTTGTTTCATATTAGAGTTAGGG		
mtAll-R23	Sequencing	ATAATGGTGCTAATTTTTGTCAGGT		
mtAll-R24	Sequencing	AGTTGGTCGTATCGGAATCG		
mtAll-R25	Sequencing	TTTGATGAGGTTTGTGGGTAGA		
mtAll-R26	Sequencing	TTCGGGGTATGGGCCCGATAGC		

Table 2.1. Continued.

Primer name	Primer function	Primer sequence	mtDNA regions (full or partial):	Size (bp)
Set 3				
mtAll-F3	PCR, Sequencing	AACTGGCTTCAATCTACTTCTCC	tRNA-Cys tRNA-Tyr Cytochrome c oxidase subunit I (*COX1*) tRNA-Ser tRNA-Asp Cytochrome c oxidase subunit II (*COX2*)	2401
mtAll-F31	Sequencing	TAATTGGAGGCTTTGGAAACTG		
mtAll-F32	Sequencing	TCCCAGTTCTAGCAGCAGGTAT		
mtAll-F33	Sequencing	CTACCCTTCATGGCGGTAA		
mtAll-F34	Sequencing	CCTGCAGGAGGAGGAGACCCAAT		
mtAll-R3	PCR, Sequencing	GGAATTGCRTCKGTTTTTAGACCT		
mtAll-R31	Sequencing	ACAAGACTAAGGAGCTAATAAGGAAAA		
mtAll-R32	Sequencing	TTAGAAGCAAATGCCTCTCAAATTAT		
mtAll-R33	Sequencing	GTCATGTAGGACAATGTCTAGTGAAG		
mtAll-R34b	Sequencing	TTGAGATTGCGATCCGTTAAT		
mtAll-R36	Sequencing	AAGCTCGAGTGTCAACGTCTATGCC		
Set 4				
mtAll-F4	PCR, Sequencing	CACTTCCAYGACCAYACCCTAATAAT	Cytochrome c oxidase subunit II (*COX2*) tRNA-Lys ATP synthase F0 subunit 8 (*ATP8*) ATP synthase F0 subunit 6 (*ATP6*) Cytochrome c oxidase subunit III (*COX3*) tRNA-Gly NADH dehydrogenase subunit 3 (*ND3*) tRNA-Arg	2758
mtAll-F41	Sequencing	CTGCGGAGCAAACCATAGC		
mtAll-F42	Sequencing	AAGAAGTCATGTGTTTTCCCTTG		
mtAll-F43	Sequencing	AAATCTCACTAGCCCATCTTCTCC		
mtAll-F45	Sequencing	TGGGCTGTCCCATCCCTAGGT		
mtAll-R4	PCR, Sequencing	TGAGTCGAAATCAYTTGTT		
mtAll-R41	Sequencing	TAAAGAAATAGGATCCTCATCAGTAAA		
mtAll-R42	Sequencing	TGGTGAGCTCAGGTAATAGTCACTC		
mtAll-R43	Sequencing	AGATAATGCTCCGGTAAGAGGTC		

Table 2.1. Continued.

Primer name	Primer function	Primer sequence	mtDNA regions (full or partial):	Size (bp)
Set 5				
mtAll-F5	PCR, Sequencing	GCCATCCAAGCTAATAATACAAGTCTA	NADH dehydrogenase subunit 3 (*ND3*) tRNA-Arg NADH dehydrogenase subunit 4L (*ND4L*) NADH dehydrogenase subunit 4 (*ND4*) tRNA-His tRNA-Ser tRNA-Leu NADH dehydrogenase subunit 5 (*ND5*)	2145
mtAll-F51	Sequencing	CTCCAATACCTACGGACTAGACTACG		
mtAll-F52	Sequencing	ACATTATATTTGAAACAACGCTTATC		
mtAll-F53	Sequencing	AATACCTCTATATGGTCTCCACCTATG		
mtAll-F54	Sequencing	TTTTAGCACGAGGCTTACAAAC		
mtAll-R5	PCR, Sequencing	AAGCAATTTGTCTTGTTGTTTTGT		
mtAll-R51a	Sequencing	CGGGTAAATGAAGGTTTAATGG		
mtAll-R52	Sequencing	TGGTAATTCGTAGTATCCCATACCC		
mtAll-R53	Sequencing	AGGAAGTATGACCCTGCGTTTA		
mtAll-R54	Sequencing	TGTTAGAGGTAAATCAGGCAAGG		
Set 6				
ND56-F1d	PCR, Sequencing	TGGTGCAACTCCARATAAAAGTAA	NADH dehydrogenase subunit 5 (*ND5*) NADH dehydrogenase subunit 6 (*ND6*) tRNA-Glu Cytochrome b (*CYTB*)	2149
ND56-3F2	Sequencing	AACAGAYGCTAACACAGCAGCTCT		
ND56-4F2	Sequencing	CCTAYATCAACACCTGAGCACTA		
ND56-45F2	Sequencing	ATTTACRGTAGCAACAGAACTTAACAA		
ND56-5F2	Sequencing	CCCCAGAATAATCCTCACG		
ND56-R1A	PCR, Sequencing	TYACATCTCGGCAAATATGG		
ND56-5R2	Sequencing	CTTGAGGGTCAAATGTTGTA		
ND56-35R2	Sequencing	GCACAGATAGCTGTRAATAAAGTAGTAAT		
ND56-45R2	Sequencing	TGTTAAGTTCTGTTGCTACYGTAAATC		

Table 2.1. Continued.

Primer name	Primer function	Primer sequence	mtDNA regions (full or partial):	Size (bp)
Set 7				
CBCR-F1d	PCR, Sequencing	ACCATGACTAATGATCTGAAAAACC	tRNA-Glu Cytochrome b (*CYTB*) tRNA-Thr tRNA-Pro D-Loop	1938
CBCR-F2	Sequencing	TTCATAGGATAYGTCCTTCC		
CBCR-F3	Sequencing	KTACGCCATYCTACGATC		
CBCR-F5	Sequencing	ATTACAATGGTCTTGTAAGCCATAAA		
CBCR-F6	Sequencing	GATAAACCATAGTCTTACATAGCACAT		
CBCR-R1d	PCR, Sequencing	CTCAGACGGCCATAGCTGA		
CBCR-R3	Sequencing	GGACRCCTCCTAGTTTGTTTGGTA		
CBCR-R4	Sequencing	CTTCCCTGAATACCCTTAGAAA		
CBCR-R5	Sequencing	GCTTTAATGTGCTATGTAAGACTATGG		
CBCR-R52	Sequencing	AAGCCTCCTCARATTCATTCTACTA		
Set 8				
mtAll-F8	PCR, Sequencing	ACCCAAARCTGANATTCT	tRNA-Pro D-Loop tRNA-Phe 12S ribosomal RNA	1731
mtAll-F82	Sequencing	GTTAATGTAGCTTAAAACAAAAGCAA		
mtAll-R8	PCR, Sequencing	TGTTACGACTTGTCTCCTCT		
mtAll-R81	Sequencing	GACTGAATTAGCAAG		
mtAll-R82	Sequencing	ACGCCGTATGCTTATTAATTTGG		

PCR amplicon sizes are based on NC_000934

Table 2.2. Proboscidean fossil calibration dates and divergence time estimates, in millions of years.

Calibration used:	Narrow Fossil[1]		Broad Fossil[2]	
	Median (95% CI)		*Median (95% CI)*	
Prior dates used to estimate divergence times				
Elephantid-Mastodon	27.0	(24-30)	27.0	(24-30)
Loxodonta-Eurasian	7.5	(6-9)	6.6	(4.2-9)
Asian-Mammoth	5.6	(4.2-7)	5.75	(3-8.5)
Forest-Savanna	Not used		Not used	
Monophyly of genera?	yes		no	
Divergence dates estimated using elephantid mitogenomes and priors only				
Loxodonta-Eurasian	6.81	(5.43-8.23)	6.06	(3.92-8.50)
Asian-Mammoth	6.01	(4.71-7.17)	5.42	(3.40-7.62)
F clade-S clade	5.51	(4.26-7.24)	4.86	(2.96-6.90)
Within-taxon coalescent dates estimated using elephantid mitogenomes and priors only				
Woolly Mammoth	1.48	(0.96-2.03)	1.33	(0.78-2.08)
Asian Elephant	1.35	(0.90-1.84)	1.20	(0.66-1.84)
African Savanna S clade	0.91	(0.60-1.30)	0.80	(0.45-1.23)
African Forest F clade	2.43	(1.68-3.38)	2.13	(1.22-3.19)
Divergence dates estimated by elephantid and mastodon mitogenomes and priors				
Elephantid-Mastodon	25.35	(23.60-26.89)	26.19	(22.48-29.24)
Loxodonta-Eurasian	7.13	(6.29-8.06)	7.43	(6.06-8.61)
Asian-Mammoth	6.35	(5.57-7.22)	6.65	(5.39-7.80)
F clade-S clade	5.83	(4.94-6.71)	6.06	(4.85-7.17)
Within-taxon coalescent dates estimated by elephantid and mastodon mitogenomes and priors				
Woolly Mammoth	1.62	(1.28-2.05)	1.70	(1.20-2.17)
Asian Elephant	1.33	(1.00-1.65)	1.41	(1.07-1.79)
African Savanna S clade	0.76	(0.57-0.99)	0.81	(0.60-1.11)
African Forest F clade	2.57	(2.07-3.15)	2.69	(1.99-3.39)
African-Eurasian ratio[3] (elephantid priors only)	0.92		0.90	
African-Eurasian ratio[4] (mastodon+elephantid priors)	0.92		0.91	
F-S clade ratio[5] (with and without mastodon)	1.06		1.25	

Footnotes:
[1]Narrow date priors are fossil dates described in Rohland et al. (2010), based on Sanders et al. (2010), with questionable taxa removed and assuming monophyly
[2]Broad date priors are fossil dates described in Rohland et al. (2010), based on Sanders et al. (2010)
[3]Ratio: F clade-S clade to Asian-Mammoth divergence, estimated using only elephantid mitogenomes and priors
[4]Ratio: F clade-S clade to Asian-Mammoth divergence, estimated using elephantid and mastodon mitogenomes and priors
[5]Ratio: F clade-S clade divergence calculated using mastodon and elephantid mitogenomes and priors vs using only elephantid mitogenomes and priors

Table 2.3. Intra-taxon mtDNA and nuclear coalescence normalized by African-Asian elephant divergence.

Taxon	mtDNA	95% CI	nuclear loci	-/+2 s.d.	mtDNA: nuclear*	mtDNA: nuclear+2s.d.
Forest elephant	36%	25-50%	30%	22-38%	1.20	0.95
Savanna elephant (S clade)	13%	9-19%	8%	4-12%	1.63	1.08
Woolly mammoth	22%	14-30%	9%	5-13%	2.44	1.69
Asian elephant	20%	13-27%	15%	9-21%	1.33	0.95

Nuclear estimates are from Rohland et al. (2010); s.d. is standard deviation
*Expected ratio is 0.25 with random progeny production and no sex differences in dispersal

CHAPTER 3: SAVANNA ELEPHANT POPULATION GENETICS

Abstract

Numerous studies have been conducted examining the phylogenetic relationship between African forest elephants and African savanna elephants. Few studies have examined the population genetic structure of African elephants. Those that have, focused on regional populations or relied heavily on mitochondrial DNA, which has been shown to be a poor indicator of population structure among African elephants. None of these studies have specifically examined the intra-species relationships among African savanna elephants. Previous studies have suggested that African savanna elephants have low genetic diversity due to a founder effect when *Elephas* went extinct in Africa followed by a recent continent wide population expansion within a short evolutionary period. Using multilocus genotype data, we extensively examined the intra-species trans-continental population genetic structure of African savanna elephants. We confirmed the relatively low genetic diversity reported by other studies, but cannot attribute this to a population bottleneck. Overall genetic substructuring was limited or absent but support for isolation by distance at the continental scale was strong. Elephants from north-central Africa appear to be somewhat distinct relative to elephants in eastern or southern Africa. Information gathered from this study should be considered by management agencies when directing efforts and resources for African savanna elephant conservation.

Introduction

Recent studies of African elephant genetics have focused heavily on the phylogenetic relationship between African forest elephants and African savanna elephants (Brandt *et al.* 2012; Debruyne 2005; Eggert *et al.* 2002; Ishida *et al.* 2011b; Johnson *et al.* 2007; Roca *et al.* 2005; Roca *et al.* 2001). The two can be distinguished from one another in the field by several features; savanna elephants have large triangular shaped ears, as well as forward and outward curved tusks, whereas forest elephants' ears are smaller and rounded, the tusks are thinner and straighter, and the overall body shape is smaller and more compact (Grubb *et al.* 2000). Genetic and morphological evidence support recognizing the two as different species (Brandt *et al.* 2012; Groves & Grubb 2000; Ishida *et al.* 2011b; Roca *et al.* 2005; Roca *et al.* 2001). Due to the treatment of all *Loxodonta* as one species, few studies have examined the genetics of savanna elephants exclusively, but in order to successfully manage and aid the recovery of this species, a solid understanding of their population dynamics is needed.

To date, no study has specifically examined the range wide, intra-species relationship of African savanna elephants. Roca and others (Roca *et al.* 2001) compared forest and savanna elephant populations using nuclear genes, noting that savanna elephants displayed fewer heterozygous sites and fewer pair-wise sequence differences than their forest counterparts. They suggest that this pattern of reduced genetic diversity indicates a recent founder event which could have occurred when *Elephas* went extinct in Africa during the Pleistocene (Kingdon 1979; Maglio 1973). Comstock and others (Comstock *et al.* 2002) examined 16 microsatellite loci in both forest and savanna elephants across their range. They too found that savanna elephants had reduced genetic diversity, indicative of a bottleneck or founder event. From their results, they also suggest that considerable gene flow has occurred between southern and eastern elephant populations. Ultimately the authors conclude that their data best supports a two species model for African elephants.

In this study we determined the extent of range wide, intra-species genetic variation for the African savanna elephant. Using multilocus genotype data we examined basic diversity indices, genetic sub-divisions, changes in effective population size and gene flow. Our findings indicate that African savanna elephants may have undergone an ancient bottleneck, have strong

support for isolation by distance at the continental scale, and limited hybridization with African forest elephants in north-central Africa.

Materials and Methods

Samples and Genotyping

This study was conducted in compliance with the University of Illinois Institutional Animal Care and Use Committed (IACUC) approved protocol number 09036. Samples were obtained in full compliance with required CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora) and other permits. Samples were collected from 558 wild African elephants (*Loxodonta*) primarily by biopsy darting (Georgiadis *et al.* 1994; Karesh *et al.* 1989; Roca *et al.* 2002) of individuals from distinct herds. DNA was isolated using kits following recommended protocols (Qiagen Inc., Valencia, CA), or standard phenol-chloroform methods (Sambrook *et al.* 1989). Eight microsatellite loci developed in savanna elephants: LAF10, LAF11, LAF12, LAF13, LAF29, LAF37, LaT05, and LaT06 (Archie *et al.* 2003; Ishida *et al.* 2011b) and 3 loci developed in Asian elephants: EMX3, EMX4, and EMX5 (Fernando *et al.* 2001), were amplified by PCR. Primers were tagged for fluorescence detection (Boutin-Ganache *et al.* 2001) and amplification followed a touchdown thermocycle profile previously described (Ishida *et al.* 2011b; Menotti-Raymond *et al.* 2005). Samples were genotyped on an ABI 3100 Genetic Analyzer and scored using GeneScan 3.7 and Genotyper 2.5 software (Applied Biosystems); alleles were subsequently binned using Allelogram (Morin *et al.* 2009). A total of 464 elephants were genotyped from 18 savanna localities including: Benoue (BE) and Waza (WA), Cameroon; Gash-Barka, Eritrea (ER); Central Kenya (KE), Mount Kenya (MK), Aberdares (AB), and Amboseli (AM), Kenya; Serengeti (SE), Ngorongoro (NG), Tanzania; Sengwa (SW), Zambezi (ZZ), and Hwange (HW), Zimbabwe; Chobe (CH), Savuti (SA), and Mashatu (MA), Botswana; Kruger (KR), South Africa; and Namibia (NA) (Table 3.1, Figure 3.1). For select analyses, an expanded dataset was used which included an additional 94 elephants genotyped for the same 11 microsatellite loci (described above) from 5 tropical forest localities including: (SL) Sierra Leone (one zoo individual); (LO) Lope, Gabon; (OD) Odzala,

Republic of Congo; (DS) Dzanga Sangha, Central African Republic; (BF) Bili Forest and 1 transitional locality (GR) Garamba, Democratic Republic of Congo (Table 3.1, Figure 3.1).

Genetic Variation Analyses

Diversity indices including number of alleles, observed heterozygosity, expected heterozygosity, linkage disequilibrium, gene diversity, genic and genotypic differentiation were calculated using a series of programs including FSTAT v. 2.9.3.2 (Goudet 1995), GENEPOP v. 4.2 (Raymond & Rousset 1995; Rousset 2008), and POPGENE v. 1.32 (Yeh & Boyle 1997). Tests for deviations from Hardy-Weinberg proportions using the Markov chain method were also performed in GENEPOP v. 4.2. F statistics, R statistics, and analysis of molecular variance (AMOVA) were calculated in ARLEQUIN v. 3.5 (Excoffier & Lischer 2010) and FSTAT. Bonferroni corrections (Rice 1989) of significance levels to compensate for multiple tests were used where appropriate.

Cluster Analyses

Population subdivisions were examined with STRUCTURE v. 2.3.3 (Hubisz *et al.* 2009). Four models were used to examine the effects of various combinations of assumptions of individual genetic ancestry and genetic relatedness among populations (Pritchard *et al.* 2000): 1) admixture with correlated allele frequencies, 2) admixture with independent allele frequencies, 3) no admixture with correlated allele frequencies and 4) no admixture with independent allele frequencies. Each model was run 3 times using values of K between 1 and 18 genetic clusters, which is the maximum number of putative populations assigned *a priori*. A second analysis was performed with STRUCTURE on the larger dataset including all 558 samples. Only the model assuming admixture with correlated allele frequencies were used as these parameters are the most biologically appropriate. The model was run 3 times using values of K between 1 and 24 genetic clusters. Each analysis was run for a minimum of 1 million Markov chain Monte Carlo steps following a burn-in of at least 100,000 steps. The uppermost hierarchical level of population structure was examined using the *ad hoc* statistic delta K based on the rate of change in ln P (D) between successive K values (Evanno *et al.* 2005), implemented in Structure Harvester (Earl & vonHoldt 2012). Multivariate representations were completed via a Factorial Correspondence Analysis (FCA) implemented in GENETIX v. 4.05 (Belkhir *et al.* 2004) and a

principal coordinate analysis (PCoA) implemented in GenALEx v. 6.5 (Peakall & Smouse 2012) using Nei's unbiased genetic distance (Nei 1978).

Gene Flow

Isolation by distance (IBD) calculations were implemented in Isolation By Distance Web Service v. 3.23 (Jensen *et al.* 2005). Pairwise geographic distances (km) were determined using the approximated center of each sampling location or country of origin. Two distances were measured, 1) assuming straight line distances between each sampling location and 2) assuming distances by way of east Africa (Aberdares, Kenya for the mid-point) for pairwise comparisons between north-central and southern localities. The second distance measure was used because pairwise straight line distances for north-central and southern localities cross the tropical forest in central Africa, which is an unlikely geographic route for gene flow thus artificially decreasing the geographic distances between these localities. Using Aberdares, Kenya as a midpoint adjusts the geographic distance to account for the tropical forest as a barrier to dispersal. Two genetic distance calculations were used: F_{ST} (Wright 1938) and Rousset's distance measure (FST/(1-FST))(Rousset 1997), and one estimate of gene flow: Slatkin's Similarity index (M=((1/FST)-1)/4) (Slatkin 1993). A Mantel test was used to assess the relationship between each genetic measure and geographic distance. Each test was performed for 30,000 iterations and compared genetic measures and geographic measures both with and without log transformations.

Evidence for recent migration was examined in BayesAss v. 3.0 (Wilson & Rannala 2003). Migration rates (posterior mean) were obtained by Markov chain Monte Carlo (MCMC) sampling from a total of 75,000,000 steps, with a discarded burn-in of at least 750,000; samples were drawn every 1,000 MCMC steps. Acceptable mixing and convergence to the stationary distribution were verified by inspection and plotting of posterior samples. The effective sample size value for each posterior mean was above 200.

Changes in N_E

To analyze for recent changes in effective population size we used the program BOTTLENECK v 1.2.02 (Cornuet & Luikart 1996). Three mutation models were considered: 1) infinite allele model (IAM), 2) step-wise mutation model (SMM), and 3) two phase model (TPM) which allows for multi-step mutations and better fits microsatellite evolution (Dirienzo *et*

al. 1994). Each mutation model was simulated for 10,000 iterations and a Wilcoxon sign-rank test was used across all loci to test for deviations from mutation-drift equilibrium, indicative of a recent reduction in effective population size.

Results

Multilocus genotypes were obtained for 464 individuals from 18 localities. Of the 11 microsatellite loci, 10 were polymorphic, and one (EMX5) was monomorphic and was excluded from subsequent analyses unless indicated otherwise. Pairwise comparisons for linkage disequilibrium were not significant ($P > 0.0009$ after correction), indicating that all loci were independent (Table 3.2). The number of alleles per locus ranged from 3 (LAF10, LAF11, EMX3, and EMX4) to 23 (LaT06); with an overall average of 9.9 alleles per locus (9.1 when EMX5 is included) across all localities (Table 3.3). Twenty private alleles were present, of which 8 occurred in north-central Africa (BE and WA, each having 4 private alleles), 6 in eastern Africa (KE, MK, AB and TA each having 1; AM having 2) and 6 in southern Africa (SW having 2; HW, CH, SA, and NA each having 1) (Figure 3.2). Fourteen of the private alleles were also found in forest or transition localities; only 1 unique allele was found in north-central Africa, 3 in eastern Africa and 2 in southern Africa. Overall expected and observed heterozygosity was 0.563 and 0.543 respectively (0.512 and 0.512 respectively when EMX5 is included). Observed heterozygosity ranged from 0.500 in ER to 0.683 in MK (0.460 in SE to 0.621 in MK when EMX5 is included) (Table 3.3). F_{IS} value for each locality was lowest in ER (-0.38) and highest in NG (0.08). At each locus F_{IS} was lowest at LAF11 (-0.08) and highest at LaT06 (0.05) (Table 3.4). Average gene diversity for all loci and localities was 0.53 (0.49 when EMX5 is included) ranging from 0.37 (ER) to 0.60 (MA) (0.37 to 0.55 for the same localities when EMX5 is included) (Table 3.5). All localities and most loci were in Hardy-Weinberg equilibrium, LAF12 was found to be heterozygote deficient (Table 3.6).

Population Differentiation

For all measures of population differentiation, including genic, genotypic, F_{ST} and R_{ST}; population differentiation was low to moderate for most pairwise comparisons (Table 3.7 and

3.8). In general, geographically distant localities had greater values of F_{ST} and R_{ST} or were significantly different for genotypic and genic differentiation than geographically closer localities. When grouped at the regional level, both genic and genotypic differentiation were significant ($P < 0.0167$ after correction) between all regional pairs (Table 3.7). For measures of F_{ST} and R_{ST}, differentiation was greatest between north-central and southern Africa ($F_{ST} = 0.06$ and $R_{ST} = 0.12$ respectively) and lowest between eastern and southern Africa ($F_{ST} = 0.01$ and $R_{ST} = 0.01$ respectively) (Table 3.8).

Sources of variation indicating differentiation were tested by analysis of molecular variance (AMOVA) and included variation among groups, among populations within groups, and within populations. Significant variation was observed in all comparisons ($P < 0.05$), within population having the highest percent variation (96.39%) and among groups having the lowest (1.36%) (Table 3.9).

Cluster Analysis

Bayesian clustering analysis was performed using STRUCTURE (Pritchard *et al.* 2000) on 464 elephants from 18 localities. No population subdivisions were apparent regardless of the model's assumptions of individual genetic ancestry and genetic relatedness (Figure 3.3). *Ad hoc* method delta-K (Evanno *et al.* 2005) useful in detecting the optimal value of K, supported 2 genetic clusters for all models (Figure 3.4, Table 3.10). It is important to note that this *ad hoc* method may not provide good support when K = 1 (Evanno *et al.* 2005). There was not an obvious trend between genetic partitions and sampling locality and this is supported by *ad hoc* method LnP(K) where K=1 (Figure 3.4, Table 3.10). Despite the lack of obvious population structure, there does appear to be a shifting pattern of cluster assignments. That is, the localities geographically farthest apart (BE and NA) show greatly differing pattern of allele assignment to each of K clusters, with a gradient of shifting allele assignments among the localities in between (Figure 3.3). This relationship persists even when the dataset is expanded to include an additional 94 individuals from 5 forest or transition localities. Bayesian cluster analysis was again performed with STRUCTURE only assuming admixture with correlated allele frequencies, with possible values for K between 1 and 24. Results confirm splitting Africa's elephants into 2 clusters (K = 2; Figure 3.6, Table 3.11), one corresponding to African forest elephants, the other to African savanna elephants as previously reported (Brandt *et al.* 2014; Ishida *et al.* 2011b).

However when 3 clusters are assumed (K=3), the gradient of allele assignments is apparent with little or no assignment of genotypes to the forest partition (Figure 3.5).

Multivariate methods were used as an alternative to Bayesian clustering analysis to visualize the genetic relationship among localities. A two-dimensional factorial correspondence analysis showed individuals grouping within their geographic region (Figure 3.7). When individuals are grouped by sampling locality, a PCoA comparing coordinates 1 and 2 clearly shows north-central localities clustering together away from eastern and southern localities. Eastern and southern localities predominately cluster by region with some overlap (Figure 3.8). To some extent this pattern is also seen in a three-dimensional FCA; however, data points are tightly clustered and the low resolution output of the software prevents this relationship from being clearly shown (data not shown).

Gene Flow

To examine African elephants for patterns of isolation by distance, pairwise genetic measures and geographic distances were examined using Mantel's test. All comparisons using F_{ST} or Roussset's distances yielded significant isolation by distance ($P < 0.05$) with regression lines of $r = 0.39$ or above (Table 3.12, Figure 3.9). Comparisons of Slatkin's similarity index with geographic distance have been reported to be useful in estimating recent isolation by distance. When populations are isolated by distance, gene flow and genetic drift are at equilibrium if given a considerable time to develop (Slatkin 1993); we therefore expect that the pairwise relationship between gene flow and geographic distance will be more negative over time. Results for this analysis were not significant ($P > 0.05$) and regression lines are close to $r = 0$ (Table 3.12, Figure 3.9), suggesting that IBD among savanna elephants may be relatively recent. Similar isolation by distance patterns were observed for both assumptions of geographic distance.

Migration was estimated for all pairs of localities and between each region. Immigration rates were low, <0.10 (Wilson & Rannala 2003) for nearly all localities with the exception of migration from KE into KR and NA into ZZ . However, migration rates were high for nearly all localities when HW was the source population (Table 3.13). This pattern was similarly observed at the regional level where migration rates were low for all areas except in north-central Africa when eastern Africa was the source population (0.27).

Changes in N_E

A two-tailed Wilcoxon sign-rank test for heterozygosity excess was performed on each locality to determine significant deviation from mutation-drift equilibrium, indicative of a recent population bottleneck. Two localities, ER and MK, were not analyzed individually due to low sample size (N = 3 for each); however, these localities were included when testing all localities combined. When the IAM of mutation was assumed, seven localities (BE, KE, AB, AM, TA, MA and NA) were significantly out of mutation-drift equilibrium (Table 3.14). For assumptions of the other mutation extreme under SMM, no individual localities were significant, though tests across all localities were (P = 0.005). Under the more realistic TPM allele mutation only one locality was significantly out of mutation-drift equilibrium, AB (P = 0.010) in eastern Africa.

Discussion

Bottleneck

By the early 20th century, the number of elephants in Africa had been reduced due to numerous factors including habitat destruction, and over hunting for meat and ivory (Douglas-Hamilton 1987). Among savanna elephants we observed low heterozygosity and low average number of alleles per locus similarly reported by other studies (Comstock *et al.* 2002; Lei *et al.* 2008). Reduced genetic diversity is expected when a population declines; when a reduction in effective population size results in future generations descending from only a few individuals (i.e. bottleneck) the number of alleles in the population will reduce faster than the heterozygosity expected at mutation-drift equilibrium. To determine if savanna elephants underwent a recent bottleneck we tested for deviations from mutation-drift equilibrium and results were significant when the stepwise mutation model was assumed (Table 3.14). The stepwise mutation model (SMM) represents an extreme in microsatellite evolution that is better suited to simple repeats (e.g. di, tri, tetra… etc.). The two-phase mutation model (TPM) primarily assumes single step changes (as the SMM does) but also allows for rare but important larger mutations (Dirienzo *et al.* 1994), which better suits our loci which also include imperfect or odd sized repeats. Under this model (TPM), deviations from equilibrium were not significant. We confirm that most

populations of African elephants have not been recently reduced in size long enough to deplete genetic diversity (Georgiadis *et al.* 1994).

While our results were not significant, deviations from mutation-drift equilibrium are only useful for determining relatively recent bottlenecks approximately $2N_e$ to $4N_e$ generations from the time the bottleneck was initiated (Cornuet & Luikart 1996). Low genetic diversity among African savanna elephants is alternatively attributed to a founder event following a bottleneck when *Elephas iolensis* went extinct in Africa at the end of the Pleistocene, allowing the range of savanna elephants to expand (Comstock *et al.* 2002; Kingdon 1979; Maglio 1973; Roca *et al.* 2001). Similarly, paleontological data supports a bottleneck early in human evolution, estimated some 2 million years ago; however, microsatellite data was not able to resolve changes in effective population size (Hawks *et al.* 2000). Our data does not support a recent bottleneck in savanna elephants, but our methods do not preclude the occurrence of a more ancient event.

Isolation by Distance

Among the 464 savanna elephants sampled from 18 localities, distinct subpopulation structuring was limited or absent. Bayesian and multivariate methods failed to consistently group individuals or sampling localities into any definitive non-overlapping population clusters. The software STRUCTURE is good for assigning individuals to one of K genetic clusters (Pritchard *et al.* 2000), among all African elephants (including both forest and savanna elephants) two genetic clusters (K=2) are clearly evident (Figure 3.5); however, no subpopulation structuring is apparent among savanna elephants (Figures 3.3). Nevertheless, there is a subtle gradient where the proportion of genotype assignments between proximate localities is more similar than between remote localities. This pattern suggests gene flow is greater between neighboring populations and limited or absent between distant populations.

Elephants are not likely to be impacted by long-term barriers to gene flow; therefore, populations may have been isolated by distance (Wright 1943). To test if savanna elephants are isolated by distance we compared three genetic measures with geographic distances (all with and without log-transformations). Methods using genetic distances (F_{ST} and Rousset's) were significant for all comparisons, strongly supporting isolation by distance in savanna elephants (Table 3.12 and Figure 3.9). Slatkin's similarity index (Slatkin 1993), which estimates pairwise gene flow, when compared to pairwise geographic distances is expected to be more negative

over time (Slatkin 1993). This measure was not significant, suggesting that the observed patterns of isolation by distance in savanna elephants may be relatively recent, less than 20,000 years (Crispo & Hendry 2005).

Estimates of migration rates between nearly all pairs of localities were less than 0.10 (Table 3.13). We did observe higher migration (between 0.10 and 0.25) rates for nearly all localities when HW (Hwange, Zimbabwe) was the source population. This pattern is likely driven by high allelic diversity in HW (Figure 3.2); recent long distance migration of elephants from HW to localities over 3000 km away is unlikely despite migration rates of nearly 0.20. Despite the capability of long distance dispersal (Thouless 1995), low migration is consistent with elephant social structure where males are the dispersing sex and females are philopatric (Archie *et al.* 2007; Hollister-Smith *et al.* 2007).

Hybridization in North-central Africa

Despite unclear population substructure among savanna elephants, the two sampling localities (BE and WA, Cameroon) in north-central Africa proved to be genetically differentiated from localities in east and southern Africa. Traditionally F_{ST} is the standard for population differentiation, and we found that pairwise comparisons of each region were low (Balloux & Lugon-Moulin 2002; Hartl & Clark 1997; Wright 1978) (Table 3.8). Contrarily, R_{ST} showed high differentiation between north-central and southern Africa, and moderate differentiation between north-central and eastern Africa (Table 3.8). As discussed previously, migration was low among savanna elephants and in this situation F_{ST} is more sensitive to polymorphisms; therefore, R_{ST} is a better indicator of subpopulation structure (Balloux & Lugon-Moulin 2002). Both multivariate methods used corroborate the R_{ST} findings that north-central localities are unique; FCA and PCoA group the northern localities having little or no overlap with eastern and southern African localities (Figure 3.7 and 3.8).

Genetic diversity among elephants in north-central localities was higher than eastern or southern Africa. Compared to other savanna elephants, those in north-central Africa had greater allele diversity. Of the 20 private alleles found 8 were from Cameroon, this is substantial considering only 6 private alleles were found in each of eastern and southern Africa despite much larger sample sizes (both in number of individuals and number of sampling localities). Additionally gene diversity per sampling locality was much higher among north-central localities

than eastern on southern. Greater genetic diversity and the presence of private alleles in Cameroon are likely due to slight hybridization with African forest elephants. Seven of the 8 private alleles in Cameroon were also genotyped in forest elephants. Hybrid individuals have previously been identified in Cameroon (Roca *et al.* 2005), and genotype assignment tests found forest elephant genetic contribution that was quite low but higher than the proportions observed in eastern and southern Africa (Ishida *et al.* 2011b).

Conservation Implications

African elephants are a flagship species in need of conservation and the results of this study should be considered for make sound management decisions. Until now, studies of savanna elephant genetics have been heavily focused on regional populations, with limited scope and geographic scale. For example, in Kenya elephants were shown to belong to three subpopulations based on mtDNA (Okello *et al.* 2008). Major motivation for research is that elephants are an important part of the eco-tourism economy (Lindsey *et al.* 2007), but human-elephant interactions with local people are not always favorable (O'Connell-Rodwell *et al.* 2000; Sukumar 1991). To resolve wildlife conflicts in Kenya, several elephants have been translocated and more relocations are planned (Okello *et al.* 2008). In other regions, elephants persist in low numbers and in complete isolation from neighboring populations. In Gash-Barka, Eritrea, fewer than 100 elephants occupy a small area of unprotected habitat; though, human-elephant conflicts are not an issue in this region, the absence of suitable habitat and low genetic diversity poses a significant risk for decline (Brandt *et al.* 2014). Wildlife management authorities in Eritrea have a much different goal than in Kenya, emphasizing the high variation in applicability of regional studies. Our data confirmed low genetic diversity among elephants across Africa. Patterns of isolation by distance at the continental scale suggest that subpopulation structure is not substantial especially across eastern and southern Africa. By contrast, the elephants of north-central Africa display greater differentiation. Consistent with previous reports, a small amount of nuclear genetic introgression was detected in savanna elephants. While low everywhere, this result of hybridization seems to have impacted the elephants of north-central Africa more than those of eastern or southern Africa.

Figures and Tables

Figure 3.1. Map of African elephant sampling locations.

This map of the continent of Africa shows the approximate locations where samples were collected. Savanna localities are indicated in black: (BE) Benoue and (WA) Waza, Cameroon; (ER) Gash-Barka, Eritrea; (KE) Central Kenya, (MK) Mount Kenya, (AB) Aberdares, and (AM) Amboseli, Kenya; (SE) Serengeti, (NG) Ngorongoro, Tanzania; (SW) Sengwa, (ZZ) Zambezi, and (HW) Hwange, Zimbabwe; (CH) Chobe, (SA) Savuti, and (MA) Mashatu, Botswana; (KR) Kruger, South Africa; and (NA) Namibia. Forest or transition localities are indicated in grey: (SL) Sierra Leone (one zoo individual); (LO) Lope, Gabon; (OD) Odzala, Republic of Congo; (DS) Dzanga Sangha, Central African Republic; (BF) Bili Forest and (GR) Garamba, Democratic Republic of Congo

Figure 3.1. Continued.

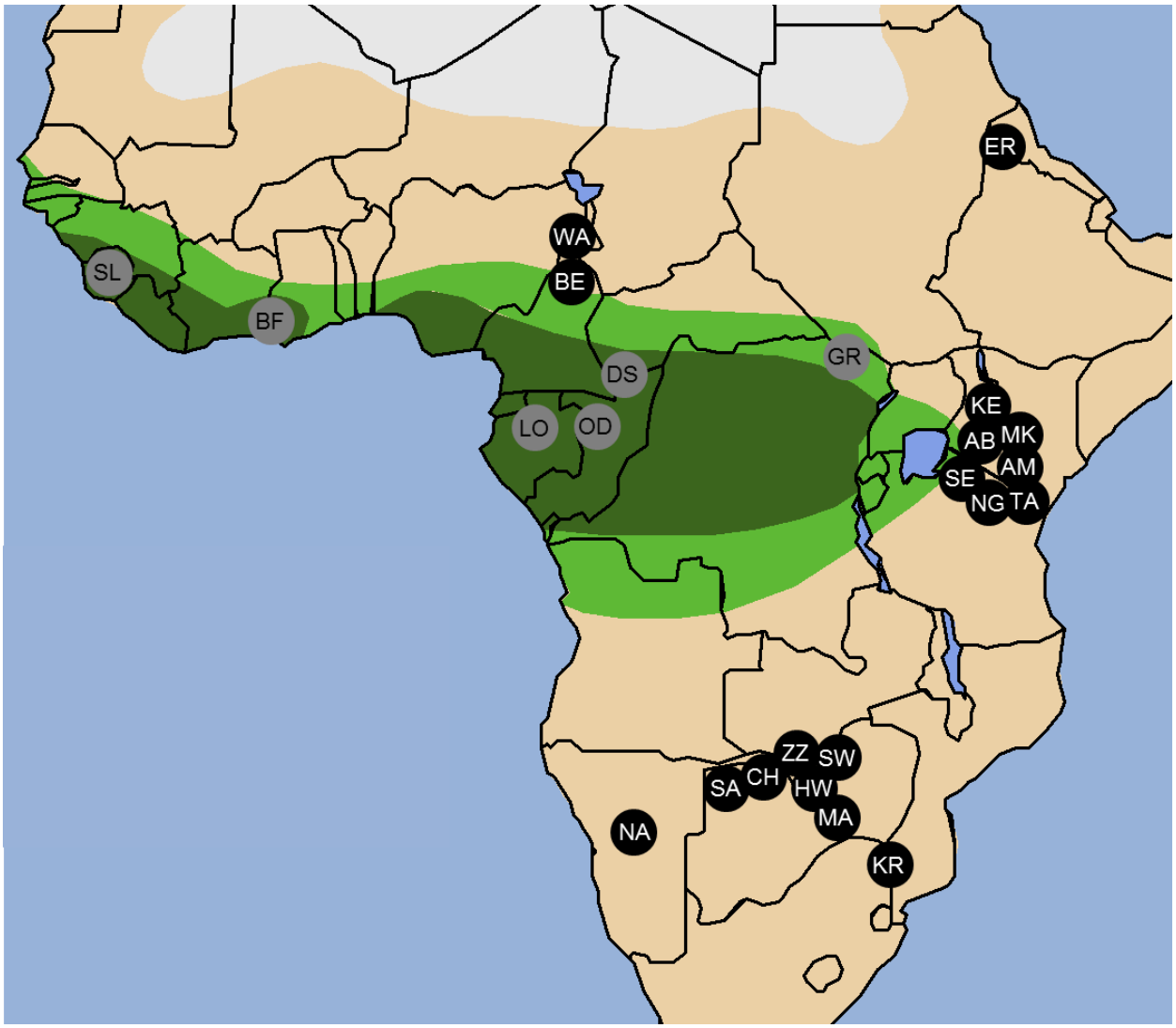


Figure 3.2. Allele frequency distribution.

Frequency distribution of 11 microsatellite loci from 18 savanna elephant localities. Horizontal axes indicates the allele size classes and the vertical axes indicate the number of alleles. Graphs in gray indicate a locus where no alleles were amplified for that locale.

Figure 3.2. Continued

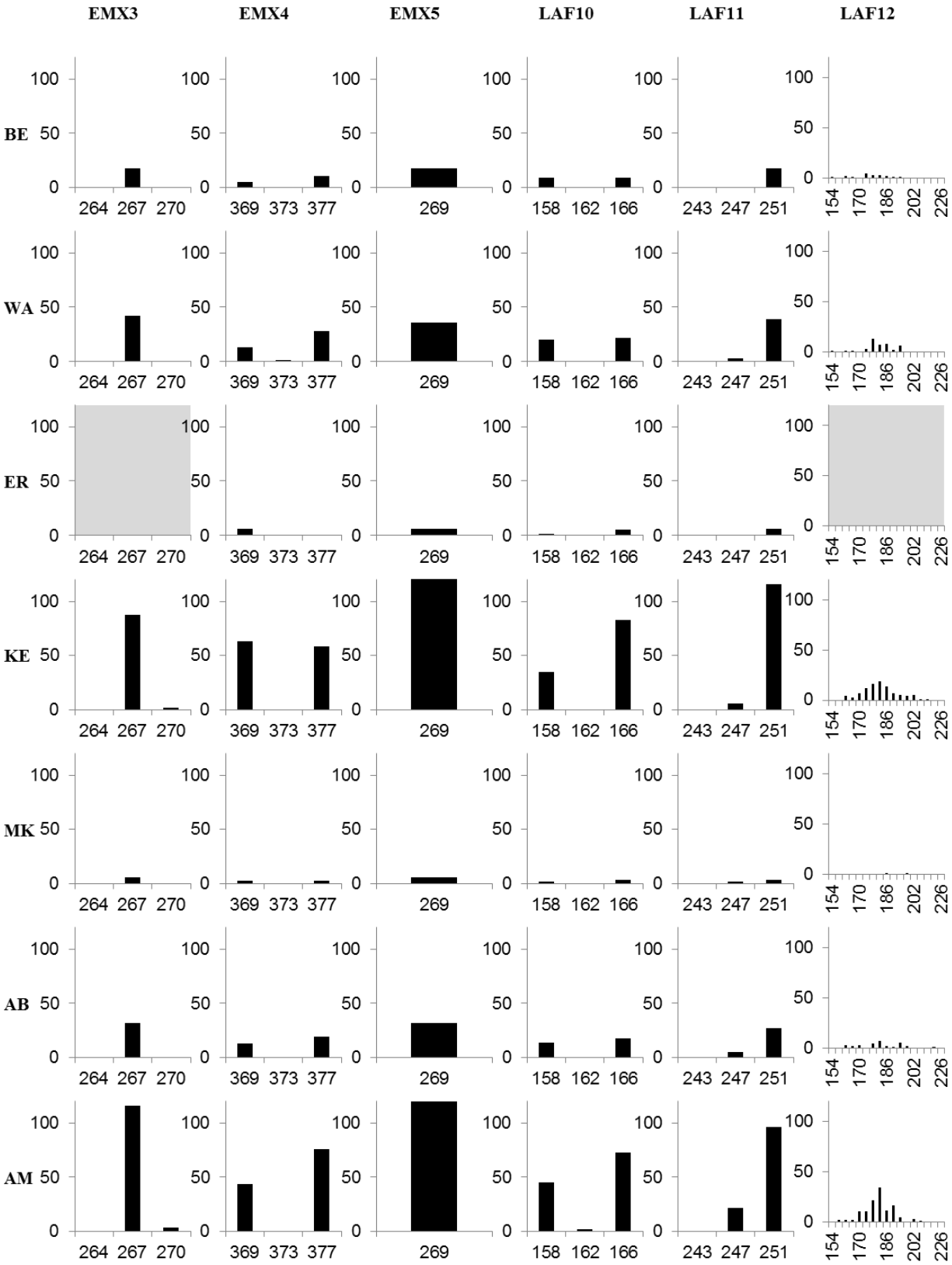


Figure 3.2. Continued.

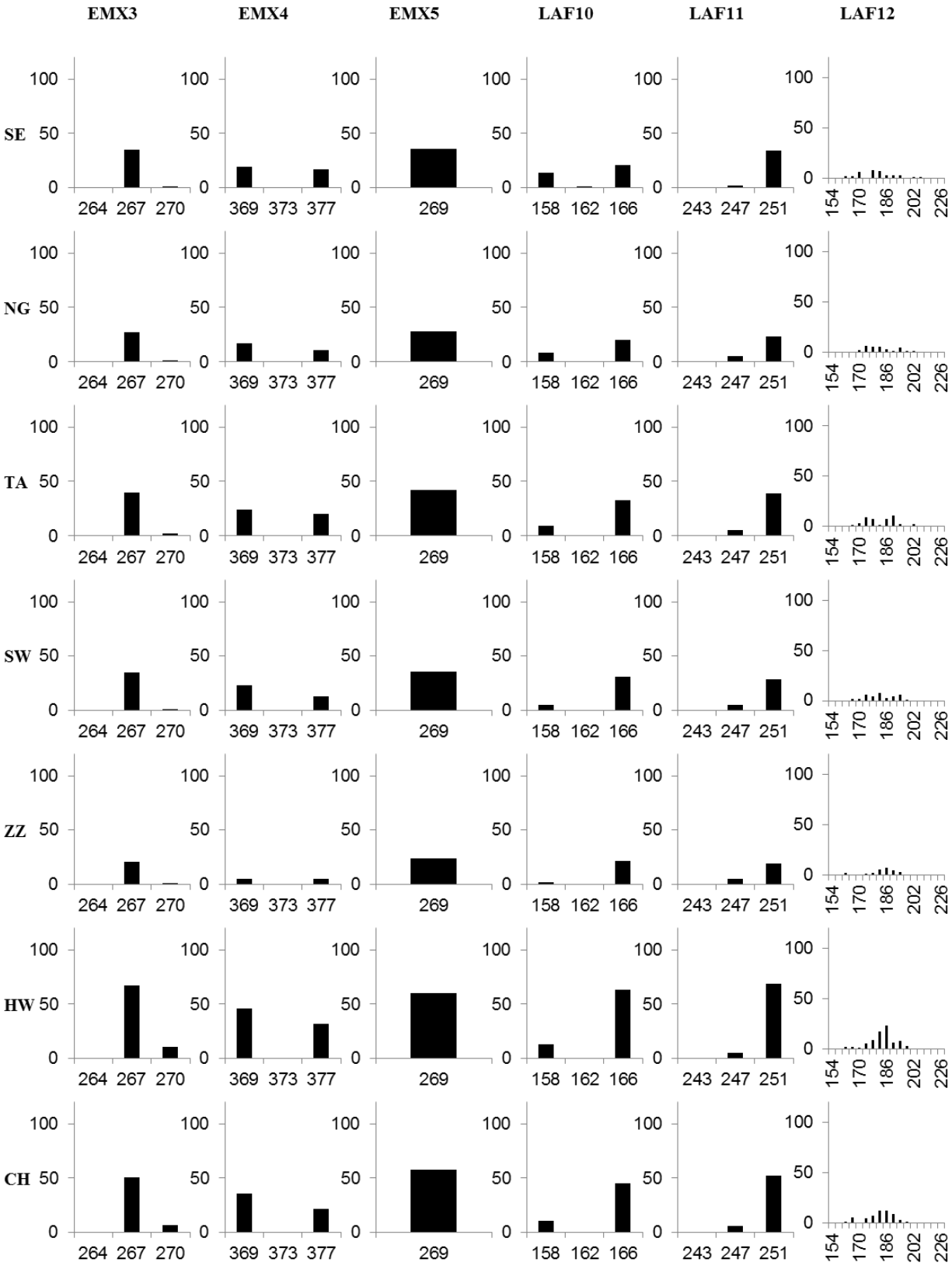


Figure 3.2. Continued.

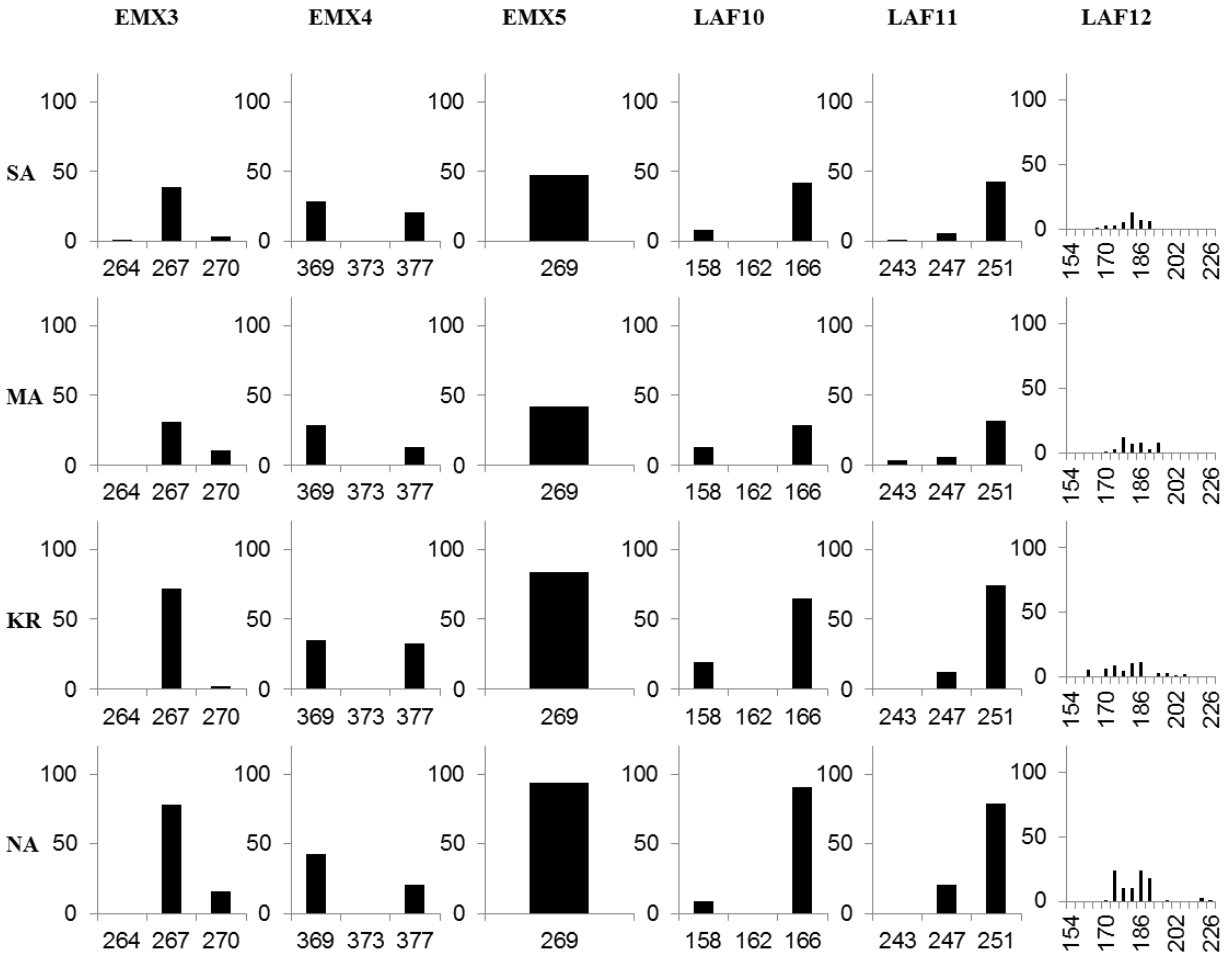


Figure 3.2. Continued

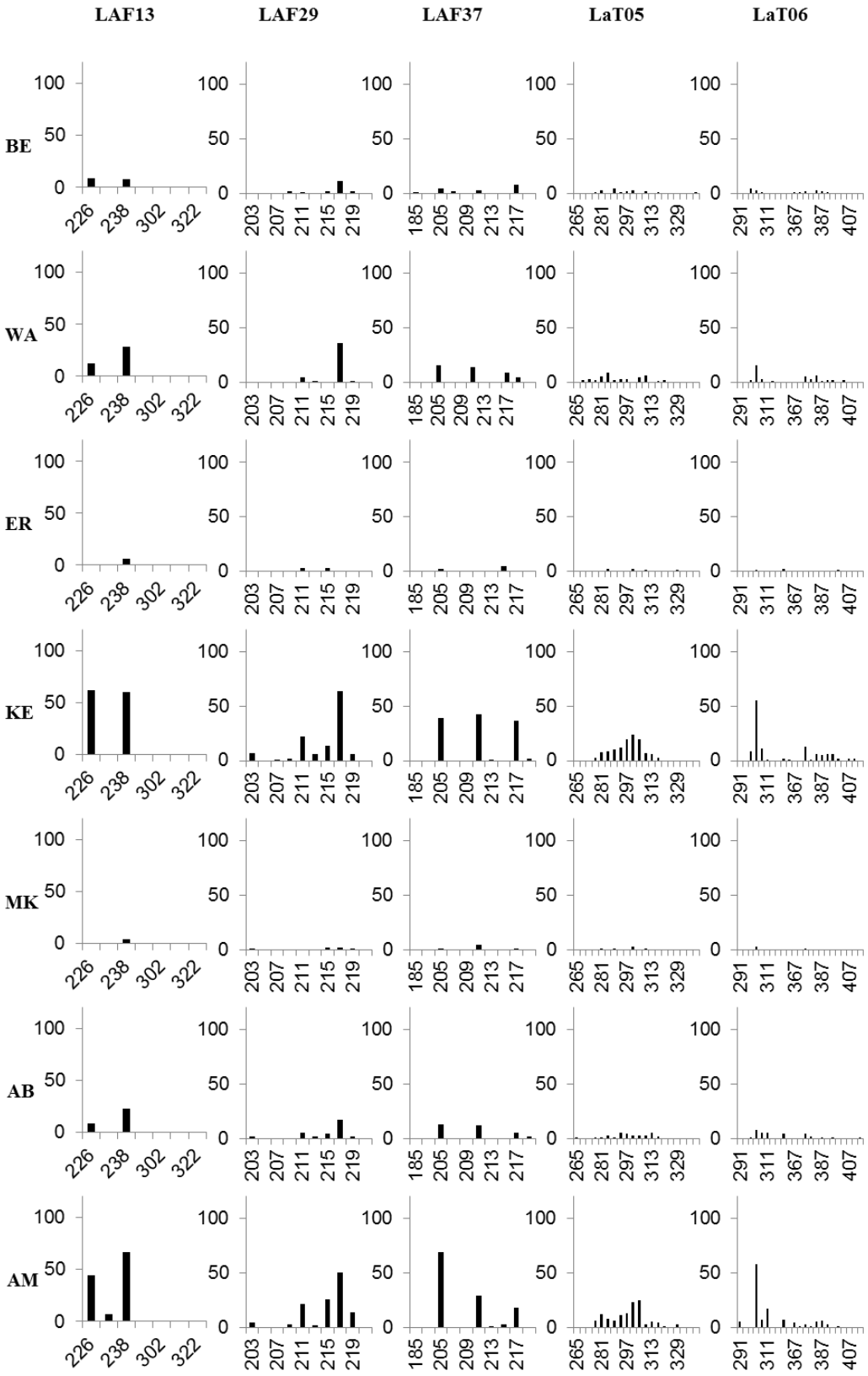


Figure 3.2. Continued

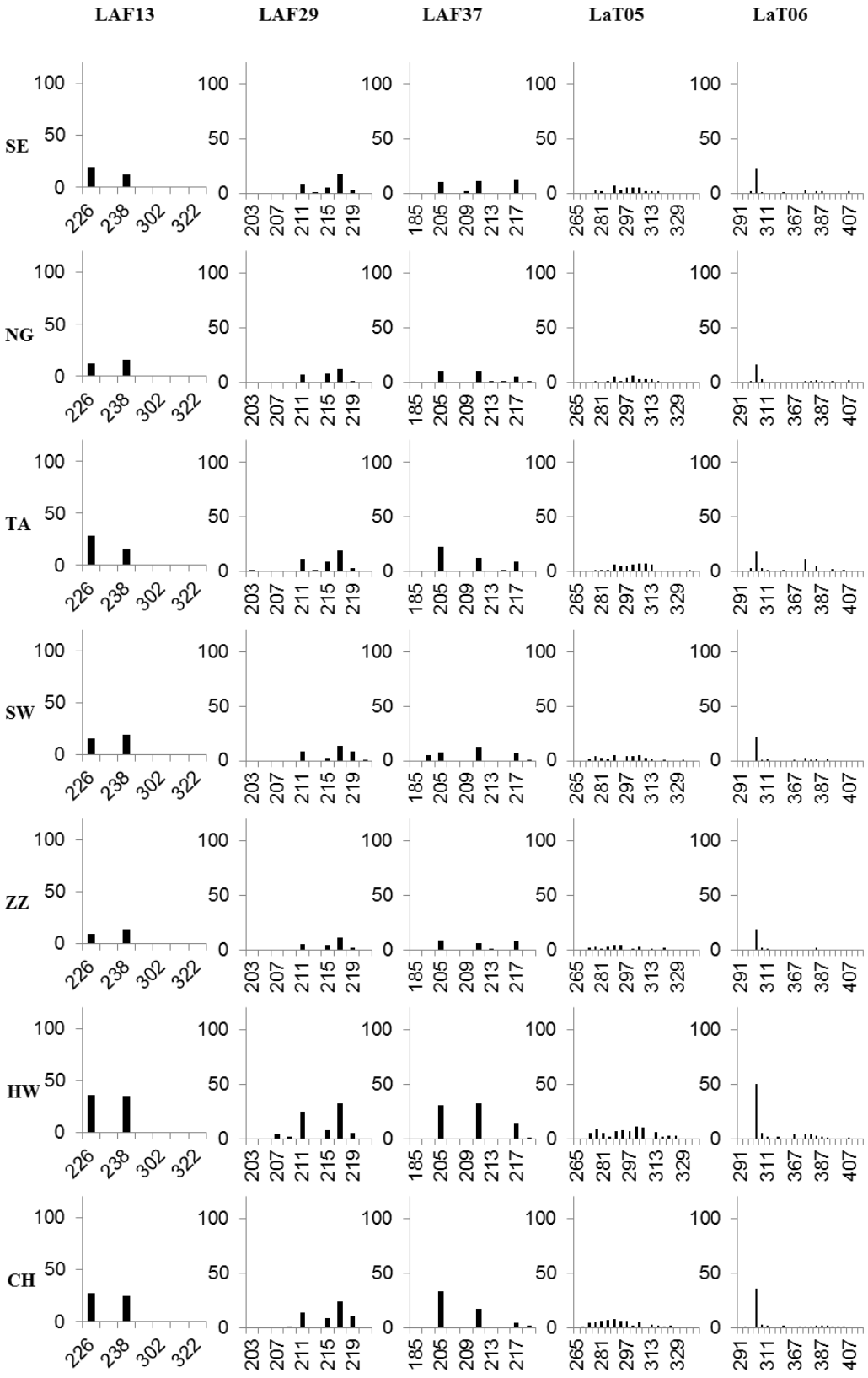


Figure 3.2. Continued

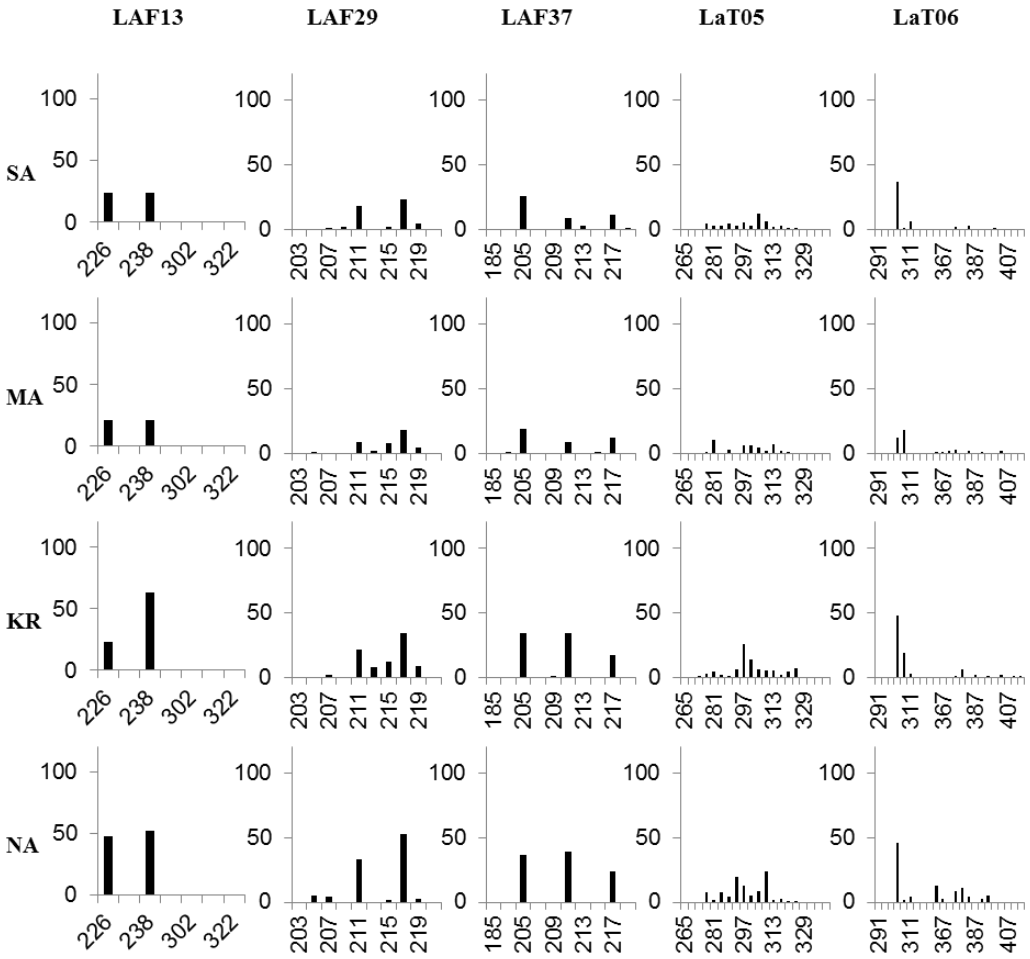


Figure 3.3. Genetic subdivision among African savanna elephants (*Loxodonta africana*). Multi-locus genotype data were used to estimate population subdivision using the program STRUCTURE (Pritchard *et al.* 2000). Four models were used to examine the effects of various combinations of assumptions of individual genetic ancestry and genetic relatedness among populations: (A) admixture with correlated allele frequencies, (B) admixture with independent allele frequencies, (C) no admixture with correlated allele frequencies and (D) no admixture with independent allele frequencies. *Ad hoc* methods support k=2; however, partitioning does not appear to follow geographic distribution. Regardless of assumptions of ancestry or relatedness, allele assignment appears to follow a gradient where the most extreme localities have a greater proportion of alleles assigning to opposite clusters, while among intermediate localities the proportion of alleles that assign to one cluster or another correlate to the geographic distance from either extreme. Abbreviations for elephant localities are: BE-Benoue, WA-Waza in Cameroon; ER-Eritrea; KE-Central Kenya/Laikipia, MK-Mount Kenya, AB-Aberdares, AM-Amboseli, in Kenya; SE-Serengeti, NG-Ngorongoro, TA-Tarangire, in Tanzania; SW-Sengwa, ZZ-Zambezi, HW-Hwange, in Zimbabwe; CH-Chobe, SA-Savuti, MA-Mashatu, in Botswana; KR-Kruger, in South Africa; NA- Namibia.

.

Figure 3.3 Continued.

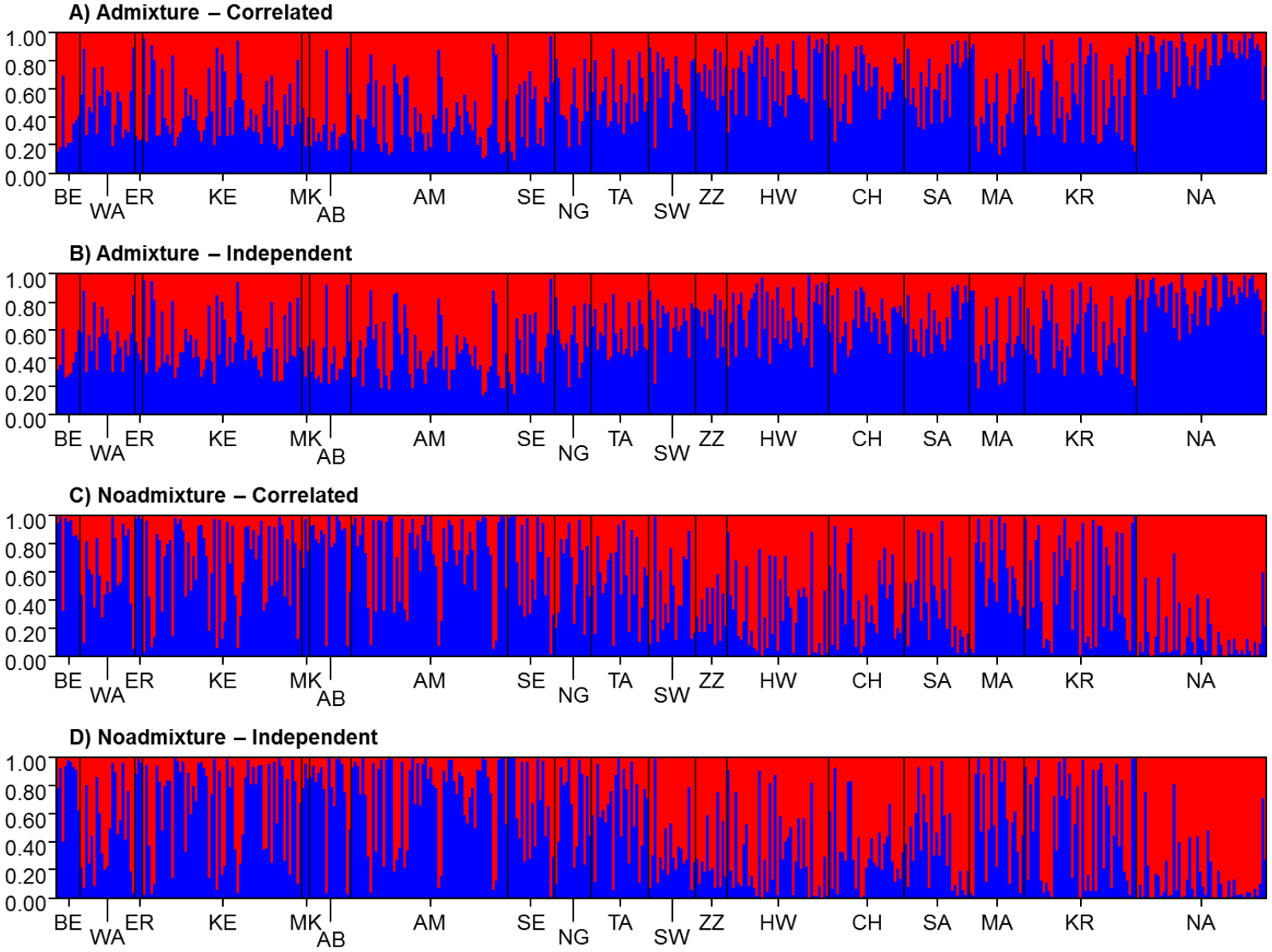


Figure 3.4. Graphical plots for examining the number of genetic subdivisions (K) in African savanna elephants (*Loxodonta africana*).

Results shown are based on the *ad hoc* method of Evanno (Evanno *et al.* 2005) as implemented in Structure Harvester (Earl & vonHoldt 2012). Combinations of assumptions of individual genetic ancestry and genetic relatedness among populations were tested: A) admixture with correlated allele frequencies, B) admixture with independent allele frequencies, C) no admixture with correlated allele frequencies and D) no admixture with independent allele frequencies. For all approaches, the Evanno method supported K = 2 clusters.

Figure 3.4. Continued.

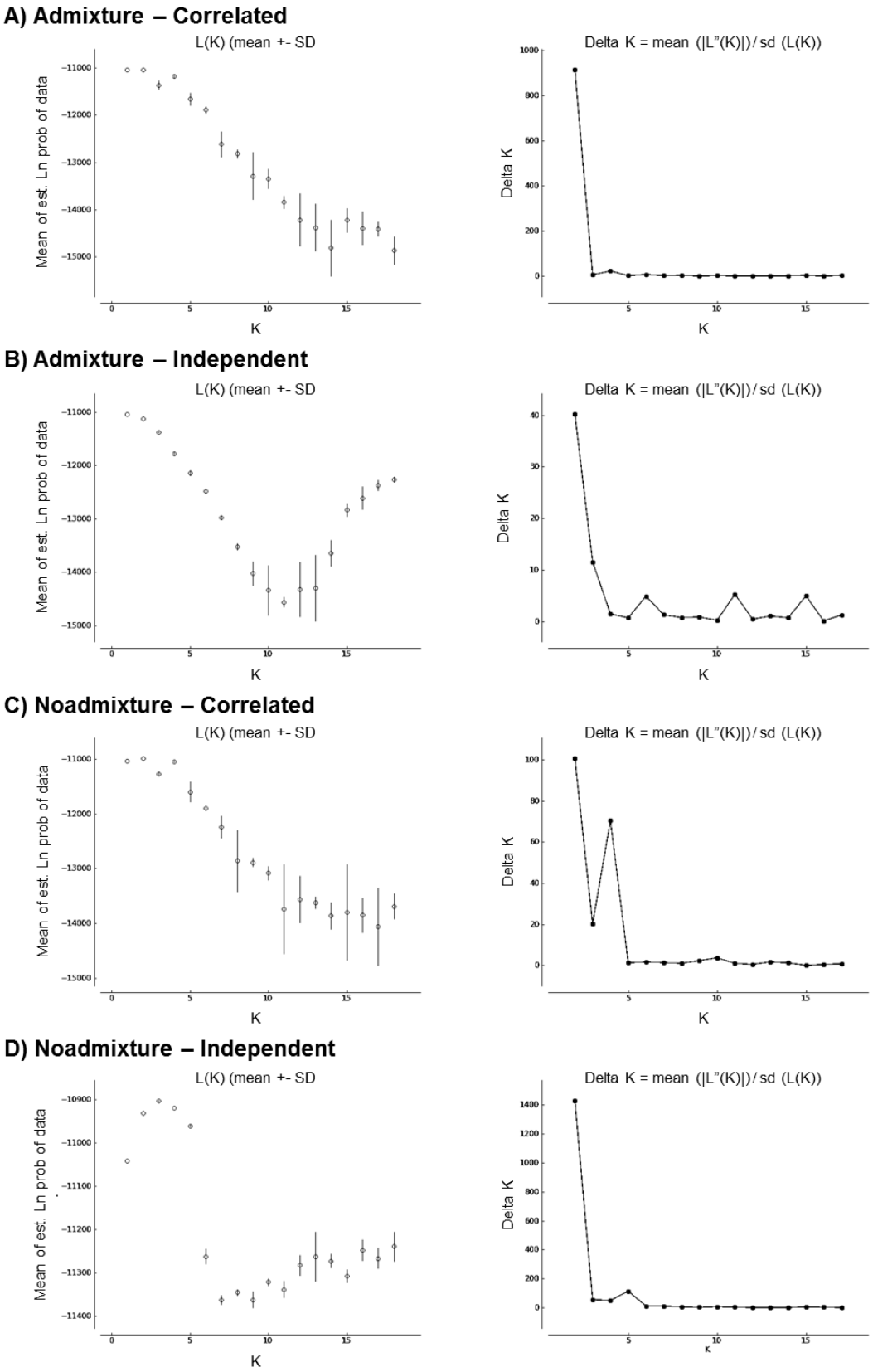


Figure 3.5. Genetic subdivision among African elephants (*Loxodonta*).

Multi-locus genotype data were used to estimate population subdivision using the program STRUCTURE (Pritchard *et al.* 2000). Two partitions that distinguish African forest elephants (green partition) from African savanna elephants (red and blue partition) were successfully reconstructed (Ishida *et al.* 2011b). No distinct clustering is apparent among individuals from savanna localities; however, individuals originating near northern Africa have a greater proportion of their genotypes assigned to the red partition whereas individuals near southern Africa have a greater proportion assigned to the blue partition. Proportion of genotype assignment to the red or blue partition is greatest in the most northern and most southern localities with those in between exhibiting genotype assignments relative to their proximity to either extreme. Abbreviations for elephant localities are: SL-Sierra Leone; LO-Lope, in Gabon; OD-Odzala, in Republic of the Congo; DS-Dzanga Sangha, in Central African Republic; BF-Bili, GR-Garamba, in the Democratic Republic of the Congo; BE-Benoue, WA-Waza in Cameroon; ER-Eritrea; KE-Central Kenya/Laikipia, MK-Mount Kenya, AB-Aberdares, AM-Amboseli, in Kenya; SE-Serengeti, NG-Ngorongoro, TA-Tarangire, in Tanzania; SW-Sengwa, ZZ-Zambezi, HW-Hwange, in Zimbabwe; CH-Chobe, SA-Savuti, MA-Mashatu, in Botswana; KR-Kruger, in South Africa; NA- Namibia.

Figure 3.5. Continued.

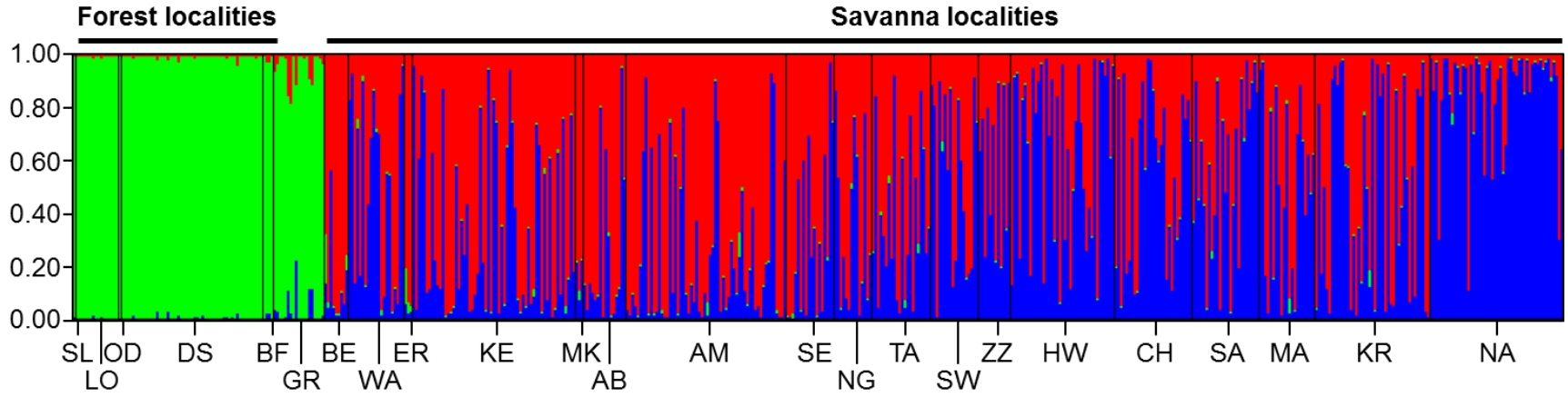


Figure 3.6. Graphical plots for examining the number of genetic subdivisions (K) in African elephants (*Loxodonta*).

Results shown are based on the *ad hoc* method of Evanno (Evanno *et al.* 2005) as implemented in Structure Harvester (Earl & vonHoldt 2012). Only the model assuming admixture with correlated allele frequencies was used. The Evanno method supports K = 2 clusters.

Figure 3.6. Continued.

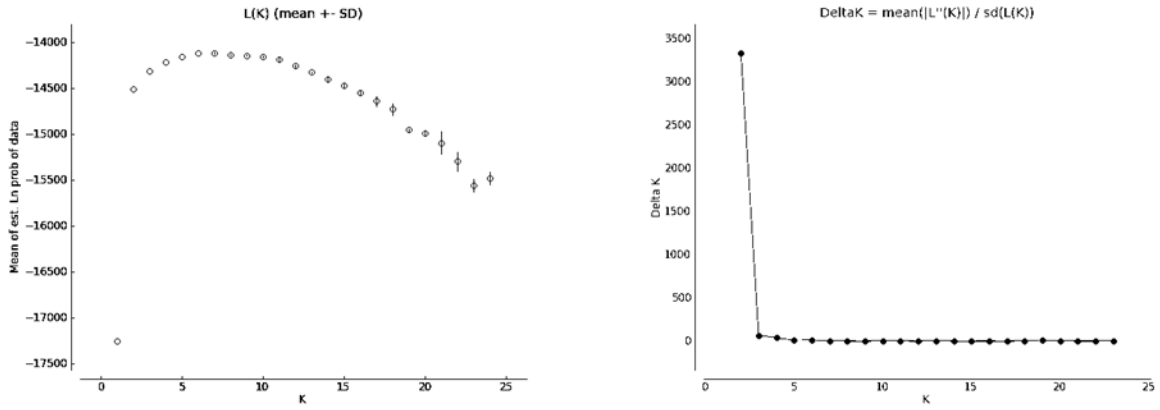


Figure 3.7. Two-dimensional factorial correspondence analysis (FCA).
This multivariate method was implemented in the program GENETIX (Belkhir *et al.* 2004). The genotypes for each individual are represented by colored squares; colors indicate origin from a locality in north-central, eastern or southern Africa (yellow, blue or white, respectively). Variation explained along axes 1 and 2 was 2.50% and 2.59%, respectively. Individuals cluster by region with eastern and southern Africa having a tighter cluster pattern than northern Africa. Some overlap among all regions is apparent, fewer individuals from northern Africa overlap with individuals from the other regions.

Figure 3.7. Continued.

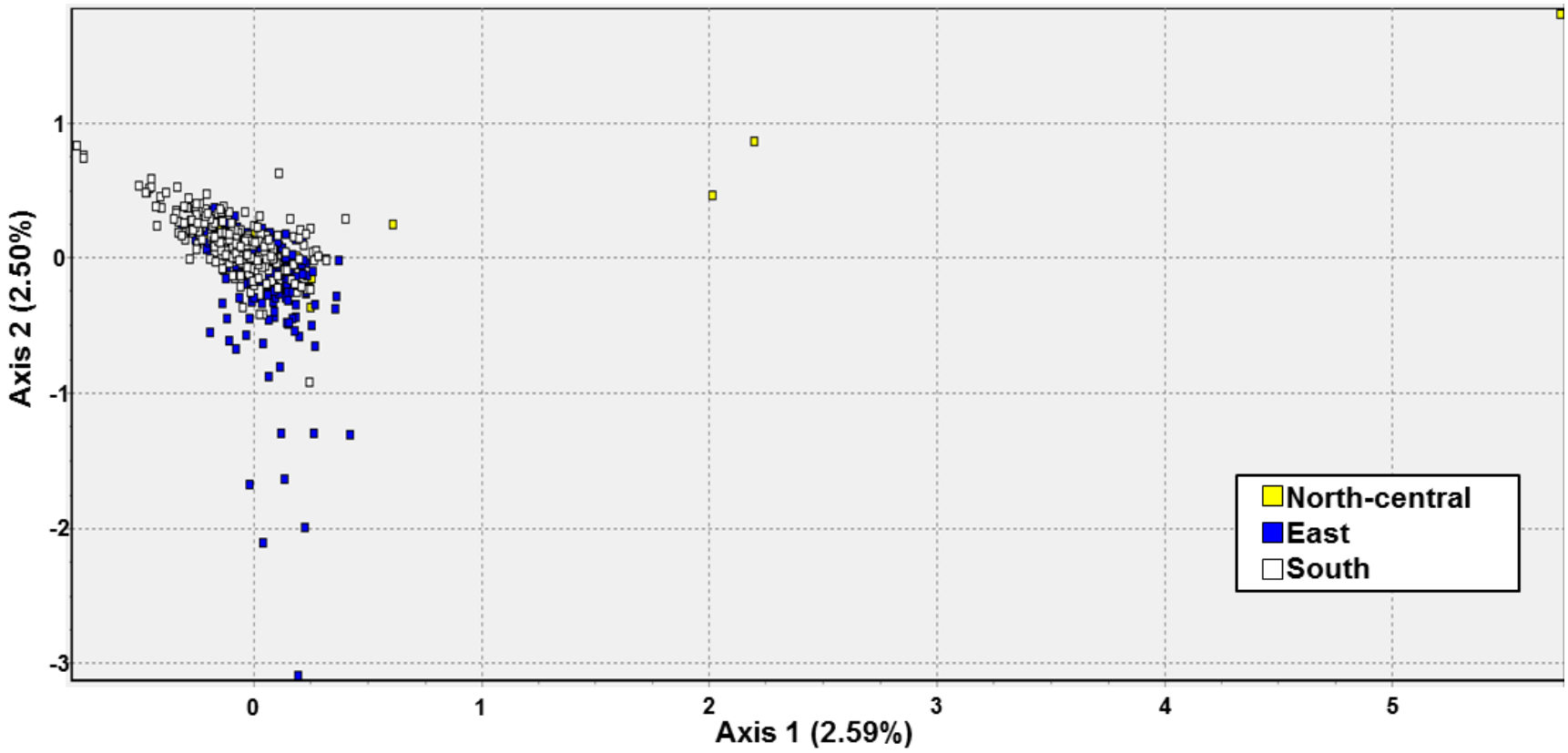


Figure 3.8. Principal Coordinate Analysis (PCoA).

This multivariate method was implemented in GenALEx (Peakall & Smouse 2012), relationships among the 18 savanna localities were estimated based on Nei's unbiased genetic distance (Nei 1978). Localities are indicated by colored diamonds; north-central (yellow), eastern (blue) and southern (white). Coordinates 1 and 2 explained the most variation; 64.57% and 15.40% respectively. Each locality groups within its respective region, some overlap is apparent between south and eastern localities. Northern African localities, BE and WA group tightly away from southern and eastern localities. Abbreviations for elephant localities are: BE-Benoue, WA-Waza in Cameroon; ER-Eritrea; KE-Central Kenya/Laikipia, MK-Mount Kenya, AB-Aberdares, AM-Amboseli, in Kenya; SE-Serengeti, NG-Ngorongoro, TA-Tarangire, in Tanzania; SW-Sengwa, ZZ-Zambezi, HW-Hwange, in Zimbabwe; CH-Chobe, SA-Savuti, MA-Mashatu, in Botswana; KR-Kruger, in South Africa; NA-Namibia.

Figure 3.8. Continued.

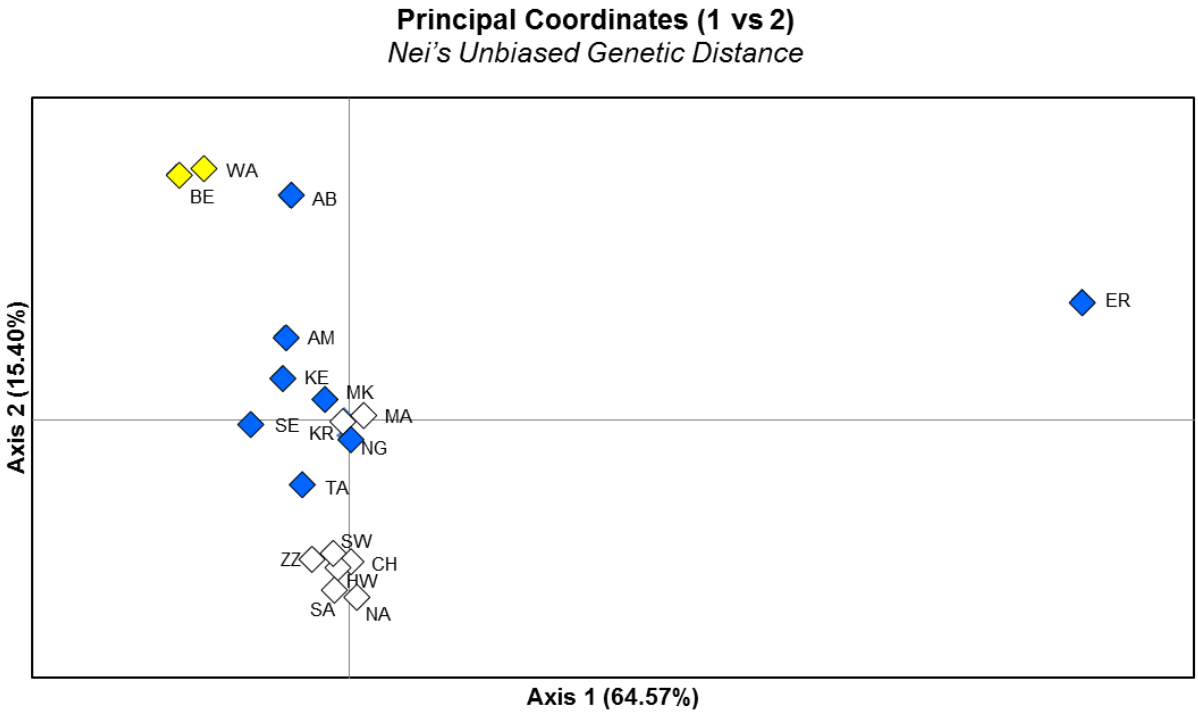


Figure 3.9. Isolation by distance.

Isolation by distance was estimated using three measures of genetic distance; F_{ST} (Wright 1938), Rousset's distance measure(FST/(1-FST))(Rousset 1997), and Slatkin's Similarity index (M=((1/FST)-1)/4) (Slatkin 1993). Genetic distances were compared to two assumptions of geographic distance: (A) straight line and (B) by way of east Africa (Aberdares, Kenya for the mid-point), as the tropical forest in central Africa would present a barrier to savanna elephant gene flow. Distance measures and log transformed distance measures were plotted in the Isolation By Distance Web Service (Jensen *et al.* 2005). Genetic distance is indicated on the vertical axes and geographic distance on the horizontal axes. All estimates using F_{ST} or Rousset's were significant with r values greater than 0.39, supporting isolation by distance among savanna elephants. Estimates using Slatkin's similarity index were not significant, suggesting patterns of isolation by distance may be recent. It is important to note that Slatkin's similarity index may not produce consistent results (Crispo & Hendry 2005).

Figure 3.9. Continued.

(A)

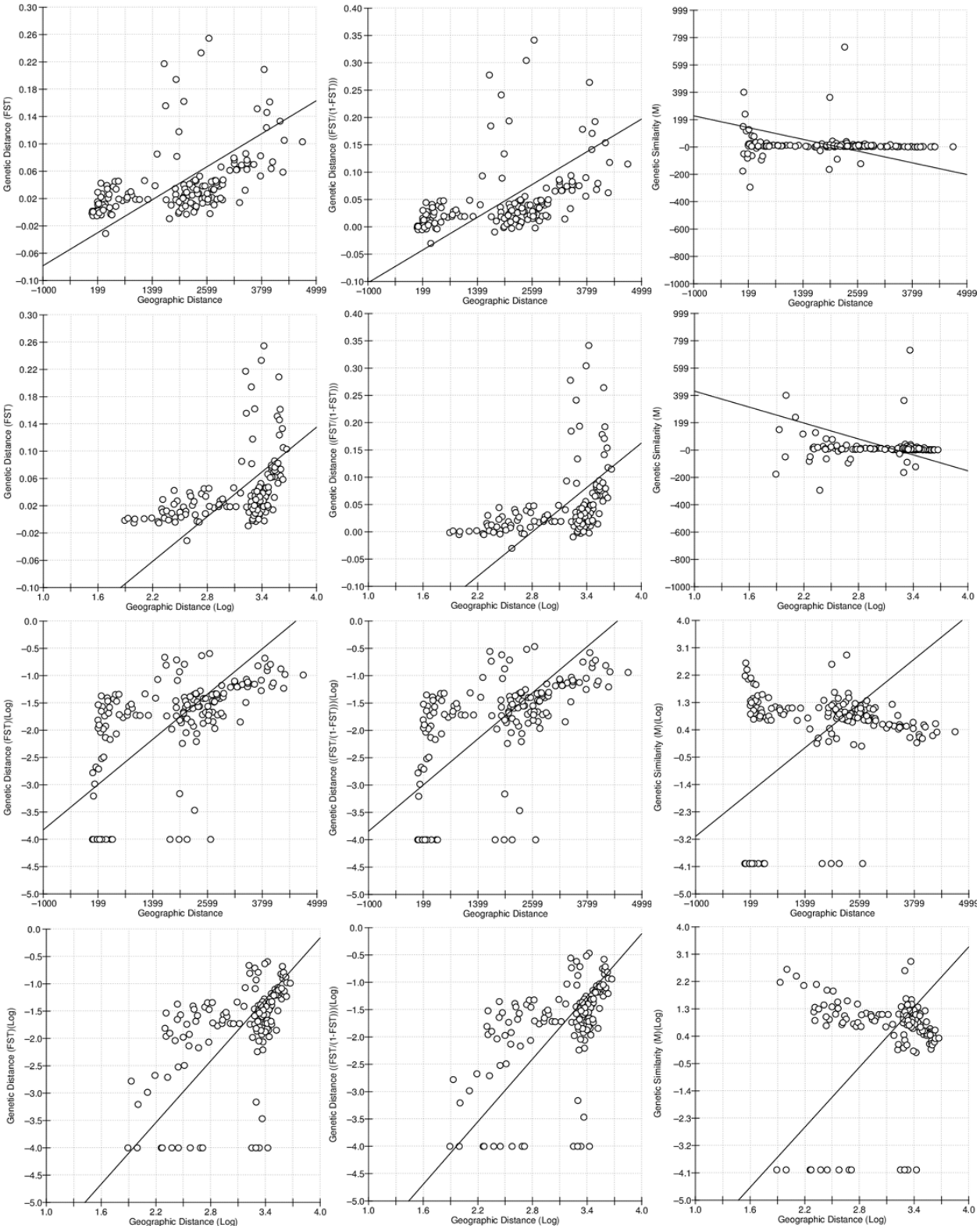


Figure 3.9. Continued.

(B)

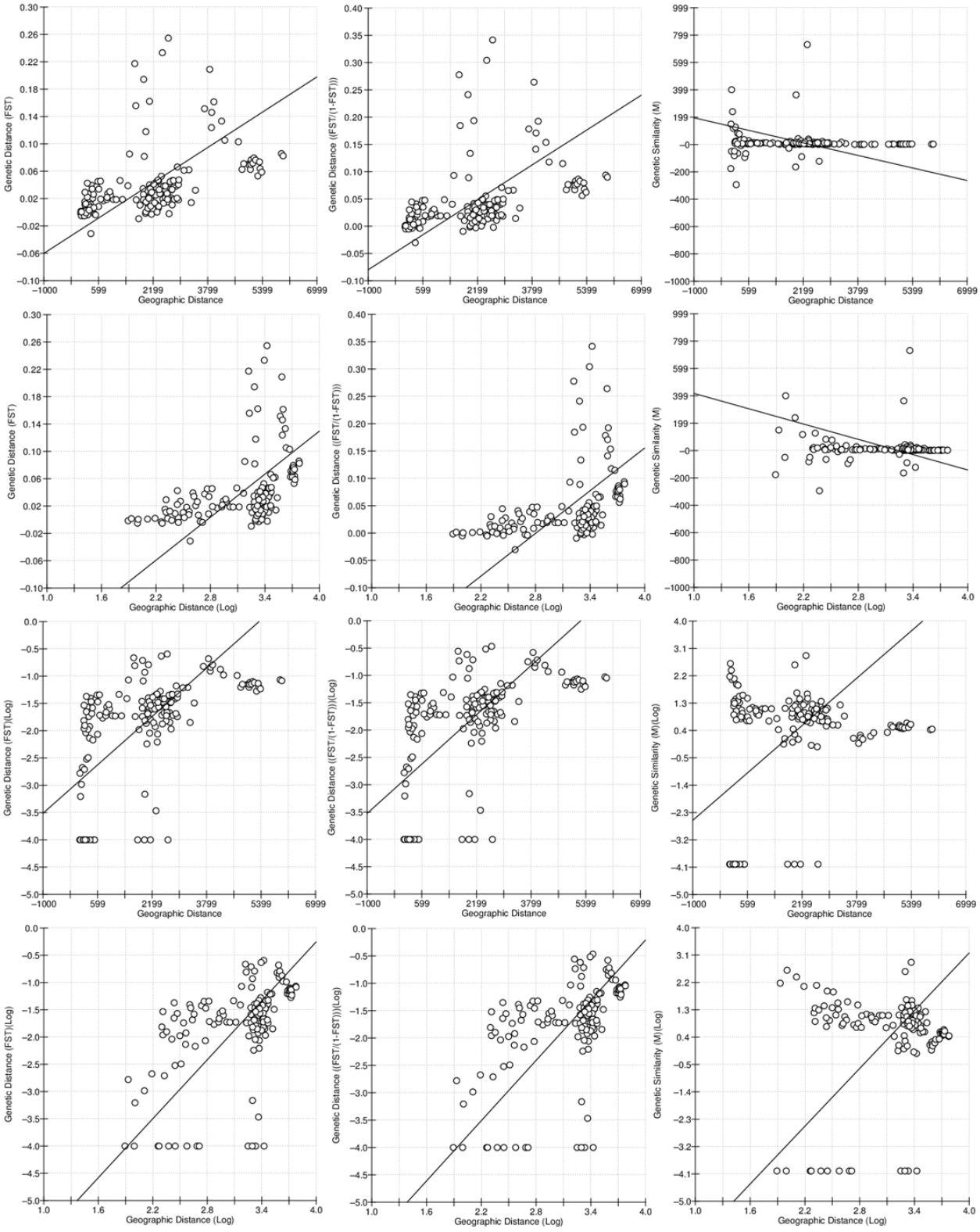


Table 3.1 Sampling location information of African elephants.

Abbreviation	Location	Country	N	Latitude*	Longitude*	Group
BE	Benoue	Cameroon	9	8.333577	13.834116	North-central
WA	Waza	Cameroon	21	11.333291	14.66656	North-central
ER	Gash-Barka	Eritrea	3	14.800638	37.218101	East
KE	Central Kenya	Kenya	61	1.241359	37.350912	East
MK	Mount Kenya	Kenya	3	-0.150225	37.317161	East
AB	Aberdares	Kenya	16	-0.416817	36.666581	East
AM	Amboseli	Kenya	60	-2.652861	37.260591	East
SE	Serengeti	Tanzania	18	-2.333957	34.832518	East
NG	Ngorongoro	Tanzania	14	-2.979159	35.446754	East
TA	Tarangire	Tanzania	22	-3.993072	35.994803	East
SW	Sengwa	Zimbabwe	18	-17.333853	28.262746	South
ZZ	Zambezi	Zimbabwe	12	-17.9784	25.709997	South
HW	Hwange	Zimbabwe	39	-18.909789	27.121291	South
CH	Chobe	Botswana	29	-18.28376	24.972847	South
SA	Savuti	Botswana	25	-18.556467	24.060078	South
MA	Mashatu	Botswana	21	-22.176508	29.046784	South
KR	Kruger	South Africa	43	-23.426041	31.337167	South
NA		Namibia	50	-22.560271	17.080684	South
SL		Sierra Leone	1	-	-	Forest
LO	Lope	Gabon	16	-	-	Forest
OD		Republic of Congo	1	-	-	Forest
DS	Dzanga Sangha	Central African Republic	53	-	-	Forest
BF	Bili Forest	Democratic Republic of Congo	4	-	-	Forest
GR	Garamba	Democratic Republic of Congo	19	-	-	Transition

*Approximated center of sampling location or country.

Table 3.2. Pairwise comparison of each locus for linkage disequilibrium.

	EMX3	EMX4	EMX5	LAF10	LAF11	LAF12	LAF13	LAF29	LAF37	LaT05	LaT06
EMX3	-	0.9964	na	0.9758	0.9742	0.9732	0.7172	0.0099	0.8839	0.9496	0.8034
EMX4		-	na	0.2416	0.3956	0.7878	0.5883	0.9925	0.9413	0.8945	0.9335
EMX5			-	na	na	na	na	na	na	na	na
LAF10				-	0.9577	0.9737	0.4437	0.6781	0.2023	0.7797	0.1369
LAF11					-	0.3474	0.5382	0.4694	0.0186	0.9375	0.1678
LAF12						-	0.7351	0.8108	0.8573	0.9914	0.6541
LAF13							-	0.8220	0.4553	0.8953	0.9059
LAF29								-	0.8929	0.9001	0.9524
LAF37									-	0.4384	0.8422
LaT05										-	0.9712
LaT06											-

Adjusted alpha-value for 5% nominal level after Bonferroni correction is 0.0009.

Table 3.3. Summary of microsatellite polymorphisms per locus for each locale.

		BE	WA	ER	KE	MK	AB	AM	SE	NG	TA	SW	ZZ	HW	CH	SA	MA	KR	NA	TOTAL
EMX3	N	9	21	0	45	3	16	60	18	14	21	18	11	39	29	22	21	37	47	431
	He	0.000	0.000	0.000	0.044	0.000	0.000	0.065	0.056	0.071	0.093	0.056	0.091	0.245	0.216	0.210	0.396	0.053	0.286	0.138
	Ho	0.000	0.000	0.000	0.044	0.000	0.000	0.067	0.056	0.071	0.095	0.056	0.091	0.282	0.241	0.227	0.429	0.054	0.298	0.139
	A	1	1	0	2	1	1	2	2	2	2	2	2	2	2	3	2	2	2	3
EMX4	N	8	21	3	61	3	16	60	18	14	22	18	5	39	29	25	21	34	32	429
	He	0.458	0.470	0.000	0.504	0.600	0.498	0.468	0.513	0.495	0.507	0.475	0.556	0.490	0.479	0.497	0.438	0.507	0.448	0.500
	Ho	0.625	0.476	0.000	0.508	1.000	0.438	0.567	0.500	0.500	0.727	0.389	0.600	0.462	0.483	0.440	0.429	0.559	0.469	0.508
	A	2	3	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3
EMX5	N	9	21	0	61	3	16	60	18	14	21	18	12	30	29	24	21	42	50	449
	He	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Ho	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	A	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
LAF10	N	9	21	3	59	3	16	60	18	14	21	18	12	38	28	25	21	42	50	458
	He	0.529	0.511	0.333	0.421	0.533	0.508	0.493	0.522	0.423	0.345	0.246	0.159	0.287	0.321	0.274	0.438	0.354	0.166	0.389
	Ho	0.333	0.381	0.333	0.492	0.667	0.375	0.467	0.444	0.286	0.333	0.278	0.000	0.290	0.321	0.240	0.429	0.405	0.180	0.354
	A	2	2	2	2	2	2	3	3	2	2	2	2	2	2	2	2	2	2	3
LAF11	N	9	21	3	61	3	16	59	18	14	22	17	12	37	29	25	21	43	50	460
	He	0.000	0.136	0.000	0.094	0.533	0.272	0.306	0.108	0.304	0.206	0.259	0.344	0.128	0.189	0.251	0.400	0.243	0.335	0.230
	Ho	0.000	0.143	0.000	0.098	0.667	0.313	0.339	0.111	0.357	0.227	0.294	0.417	0.081	0.207	0.280	0.429	0.279	0.340	0.244
	A	1	2	1	2	2	2	2	2	2	2	2	2	2	2	3	3	2	2	3
LAF12	N	9	21	0	49	1	15	58	18	14	21	18	12	38	27	19	21	27	46	414
	He	0.909	0.830	0.000	0.890	1.000	0.894	0.843	0.881	0.881	0.851	0.881	0.848	0.831	0.855	0.815	0.827	0.878	0.809	0.869
	Ho	0.889	0.810	0.000	0.837	1.000	0.800	0.793	0.833	0.857	0.952	0.889	0.833	0.737	0.815	0.842	0.714	0.852	0.739	0.812
	A	9	9	0	13	2	10	12	10	9	9	9	7	10	9	7	7	10	9	17
LAF13	N	9	21	3	61	3	16	59	16	14	22	18	12	36	26	24	21	43	50	454
	He	0.582	0.484	0.000	0.504	0.600	0.417	0.540	0.522	0.508	0.474	0.538	0.507	0.521	0.509	0.511	0.512	0.396	0.504	0.511
	Ho	0.778	0.571	0.000	0.525	0.667	0.438	0.525	0.500	0.571	0.455	0.500	0.667	0.500	0.654	0.500	0.619	0.442	0.440	0.518
	A	3	4	1	2	3	2	3	3	2	2	3	2	3	2	2	2	2	2	7
LAF29	N	9	21	3	61	3	16	60	18	14	22	18	11	38	29	25	21	43	50	462
	He	0.621	0.261	0.600	0.676	0.867	0.688	0.739	0.679	0.696	0.720	0.737	0.688	0.705	0.729	0.662	0.740	0.753	0.611	0.695
	Ho	0.333	0.286	1.000	0.705	1.000	0.813	0.733	0.500	0.643	0.636	0.722	0.636	0.632	0.862	0.600	0.619	0.744	0.620	0.665
	A	5	4	2	8	4	6	7	5	4	6	5	4	6	5	6	6	6	6	10

Table 3.3. Continued.

		BE	WA	ER	KE	MK	AB	AM	SE	NG	TA	SW	ZZ	HW	CH	SA	MA	KR	NA	TOTAL
LAF37	N	9	21	3	61	3	16	60	18	14	22	17	12	39	28	25	21	43	50	462
	He	0.752	0.724	0.533	0.687	0.600	0.688	0.593	0.716	0.735	0.648	0.756	0.714	0.650	0.564	0.658	0.683	0.656	0.660	0.678
	Ho	0.778	0.714	0.667	0.656	0.667	0.563	0.550	0.667	0.571	0.591	0.706	0.750	0.744	0.750	0.600	0.714	0.674	0.600	0.652
	A	5	4	2	5	3	4	5	4	6	4	5	4	4	4	5	5	4	3	10
LaT05	N	9	21	3	61	3	16	60	18	14	22	18	12	39	29	25	21	43	50	464
	He	0.909	0.907	0.867	0.882	0.800	0.921	0.884	0.903	0.894	0.895	0.925	0.917	0.917	0.924	0.903	0.876	0.862	0.867	0.909
	Ho	1.000	0.952	1.000	0.803	0.667	0.938	0.900	0.944	0.786	0.955	1.000	0.833	0.974	0.931	0.920	0.857	0.837	0.820	0.888
	A	9	12	4	11	4	12	13	10	10	11	12	10	13	14	13	10	14	13	20
LaT06	N	9	21	2	61	2	16	59	18	14	22	17	12	39	28	25	21	42	50	458
	He	0.909	0.837	0.833	0.768	0.500	0.877	0.728	0.587	0.669	0.766	0.578	0.373	0.581	0.586	0.441	0.739	0.622	0.751	0.714
	Ho	0.889	0.810	1.000	0.672	0.500	0.875	0.729	0.500	0.571	0.818	0.529	0.333	0.564	0.571	0.480	0.667	0.595	0.680	0.649
	A	9	11	3	15	2	10	13	8	9	9	8	4	11	14	6	9	10	10	23
ALL	N	9	21	3	61	3	16	60	18	14	22	18	12	39	29	25	21	43	50	464
	He	0.515	0.469	0.396	0.497	0.549	0.524	0.514	0.499	0.516	0.501	0.495	0.473	0.487	0.488	0.475	0.550	0.484	0.494	0.512
	Ho	0.511	0.468	0.500	0.485	0.621	0.505	0.515	0.460	0.474	0.526	0.488	0.469	0.479	0.531	0.466	0.537	0.495	0.471	0.512
	A	47	53	16	63	26	52	63	50	49	50	51	40	56	57	50	49	55	52	100
ALL*	N	9	21	3	61	3	16	60	18	14	22	18	12	39	29	25	21	43	50	464
	He	0.567	0.516	0.396	0.547	0.603	0.576	0.566	0.549	0.568	0.551	0.545	0.520	0.535	0.537	0.522	0.605	0.533	0.544	0.563
	Ho	0.563	0.514	0.500	0.534	0.683	0.555	0.567	0.506	0.521	0.579	0.536	0.516	0.527	0.584	0.513	0.591	0.544	0.519	0.543
	A	46	52	16	62	25	51	62	49	48	49	50	39	55	56	49	48	54	51	99

Variables are number of individuals genotyped (N), expected and observed heterozygosity (H_E calculated using Levene's (Levene 1949) and H_O respectively), and total number of alleles (A) per locus. *Calculations across all loci when EMX5 is excluded.

Table 3.4. Estimation of the inbreeding coefficient (F_{IS}) per locus for each locale and overall.

	BE	WA	ER	KE	MK	AB	AM	SE	NG	TA	SW	ZZ	HW	CH	SA	MA	KR	NA	ALL
EMX3	-	-	-	-0.01	-	-	-0.03	0.00	0.00	-0.03	0.00	0.00	-0.15	-0.12	-0.08	-0.08	-0.01	-0.04	-0.07
EMX4	-0.40	-0.01	-	-0.01	-1.00	0.13	-0.21	0.03	-0.01	-0.45	0.18	-0.09	0.06	-0.01	0.12	0.02	-0.10	-0.05	-0.06
EMX5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LAF10	0.38	0.26	0.00	-0.17	-0.33	0.27	0.05	0.15	0.33	0.03	-0.13	1.00	-0.01	0.00	0.13	0.02	-0.14	-0.09	0.04
LAF11	-	-0.05	-	-0.04	-0.33	-0.15	-0.11	-0.03	-0.18	-0.11	-0.14	-0.22	0.37	-0.10	-0.12	-0.07	-0.15	-0.01	-0.08
LAF12	0.02	0.03	-	0.06	-1.00	0.11	0.06	0.06	0.03	-0.12	-0.01	0.02	0.11	0.05	-0.03	0.14	0.03	0.09	0.05
LAF13	-0.37	-0.19	-	-0.04	-0.14	-0.05	0.03	0.04	-0.13	0.04	0.07	-0.33	0.04	-0.29	0.02	-0.22	-0.12	0.13	-0.04
LAF29	0.48	-0.10	-1.00	-0.04	-0.20	-0.19	0.01	0.27	0.08	0.12	0.02	0.08	0.11	-0.19	0.10	0.17	0.01	-0.02	0.02
LAF37	-0.04	0.01	-0.33	0.05	-0.14	0.19	0.07	0.07	0.23	0.09	0.07	-0.05	-0.15	-0.34	0.09	-0.05	-0.03	0.09	0.02
LaT05	-0.11	-0.05	-0.20	0.09	0.20	-0.02	-0.02	-0.05	0.13	-0.07	-0.08	0.09	-0.06	-0.01	-0.02	0.02	0.03	0.06	0.00
LaT06	0.02	0.03	-0.33	0.13	0.00	0.00	0.00	0.15	0.15	-0.07	0.09	0.11	0.03	0.02	-0.09	0.10	0.04	0.10	0.05
ALL	0.01	0.00	-0.38	0.02	-0.21	0.04	0.00	0.08	0.08	-0.05	0.02	0.01	0.02	-0.09	0.02	0.02	-0.02	0.05	

Table 3.5. Gene diversity per locus for each locale.

	BE	WA	ER	KE	MK	AB	AM	SE	NG	TA	SW	ZZ	HW	CH	SA	MA	KR	NA	ALL
EMX3	0.00	0.00	-	0.04	0.00	0.00	0.07	0.06	0.07	0.09	0.06	0.09	0.25	0.22	0.21	0.40	0.05	0.81	0.14
EMX4	0.45	0.47	0.00	0.50	0.50	0.50	0.47	0.51	0.50	0.50	0.48	0.55	0.49	0.48	0.50	0.44	0.51	0.45	0.46
EMX5	0.00	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LAF10	0.54	0.51	0.33	0.42	0.50	0.51	0.49	0.53	0.43	0.35	0.25	0.17	0.29	0.32	0.28	0.44	0.35	0.17	0.38
LAF11	0.00	0.14	0.00	0.09	0.50	0.27	0.31	0.11	0.30	0.21	0.26	0.34	0.13	0.19	0.25	0.40	0.24	0.34	0.23
LAF12	0.91	0.83	-	0.89	-	0.90	0.84	0.88	0.88	0.85	0.88	0.85	0.83	0.86	0.81	0.83	0.88	0.17	0.82
LAF13	0.57	0.48	0.00	0.50	0.58	0.42	0.54	0.52	0.51	0.47	0.54	0.50	0.52	0.51	0.51	0.51	0.40	0.51	0.48
LAF29	0.64	0.26	0.50	0.68	0.83	0.68	0.74	0.69	0.70	0.72	0.74	0.69	0.71	0.73	0.66	0.74	0.75	0.29	0.65
LAF37	0.75	0.72	0.50	0.69	0.58	0.69	0.59	0.72	0.74	0.65	0.76	0.71	0.65	0.56	0.66	0.68	0.66	0.61	0.66
LaT05	0.90	0.91	0.83	0.88	0.83	0.92	0.88	0.90	0.90	0.89	0.92	0.92	0.92	0.92	0.90	0.88	0.86	0.87	0.89
LaT06	0.91	0.84	0.75	0.77	0.50	0.88	0.73	0.59	0.67	0.77	0.58	0.38	0.58	0.59	0.44	0.74	0.62	0.75	0.67
ALL	0.52	0.47	0.36	0.50	0.48	0.52	0.51	0.50	0.52	0.50	0.50	0.47	0.49	0.49	0.47	0.55	0.48	0.45	0.49
ALL*	0.56	0.51	0.36	0.54	0.53	0.57	0.56	0.55	0.56	0.55	0.54	0.52	0.53	0.53	0.52	0.60	0.53	0.49	0.53

*Calculation across all loci when EMX5 is excluded.

Table 3.6. Hardy Weinberg exact test.

Locale	P-value	Locus	P-value
BE	0.27	EMX3	0.93
WA	0.22	EMX4	0.87
ER	1.00	EMX5	-
KE	0.10	LAF10	0.22
MK	0.93	LAF11	0.92
AB	0.42	LAF12	**0.02**
AM	0.46	LAF13	0.85
SE	0.09	LAF29	0.45
NG	0.06	LAF37	0.16
TA	0.92	LaT05	0.59
SW	0.39	LaT06	0.12
ZZ	0.15		
HW	0.43		
CH	0.98		
SA	0.61		
MA	0.37		
KR	0.40		
NA	0.30		

Values in bold are significant ($P < 0.05$), indicating heterozygote deficiency.

Table 3.7. Genic and Genotypic differentiation between each locale and region pair.

	BE	WA	ER	KE	MK	AB	AM	SE	NG	TA	SW	ZZ	HW	CH	SA	MA	KR	NA
BE	--	0.0120	<0.0001	0.0103	0.4404	0.0020	<0.0001	0.1938	0.0263	<0.0001	0.0003	0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
WA	0.0219	--	<0.0001	<0.0001	0.0103	0.0102	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0003	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	< 0.0001
ER	<0.0001	<0.0001	--	<0.0001	0.1091	0.0001	<0.0001	<0.0001	0.0019	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	< 0.0001
KE	0.0151	<0.0001	<0.0001	--	0.4593	0.0049	<0.0001	0.5368	0.7549	0.1200	0.0001	0.0012	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	< 0.0001
MK	0.4370	0.0158	0.1278	0.4501	--	0.8143	0.2296	0.2991	0.8227	0.4144	0.7011	0.2436	0.1408	0.0508	0.0798	0.2157	0.3721	0.0035
AB	0.0049	0.0075	0.0008	0.0079	0.8808	--	0.0037	0.0175	0.1913	0.0001	0.0006	0.0009	<0.0001	<0.0001	<0.0001	<0.0001	0.0021	< 0.0001
AM	<0.0001	<0.0001	<0.0001	<0.0001	0.3299	0.0046	--	0.0006	0.0272	<0.0001	<0.0001	0.0100	-	<0.0001	<0.0001	<0.0001	<0.0001	< 0.0001
SE	0.2763	<0.0001	<0.0001	0.5497	0.3948	0.0263	0.0016	--	0.5469	0.0901	0.0406	0.0372	0.0023	0.0002	0.0044	<0.0001	<0.0001	< 0.0001
NG	0.0478	0.0002	0.0033	0.7650	0.7918	0.2546	0.0285	0.6501	--	0.6508	0.6562	0.4728	0.1303	0.0832	0.0942	0.0508	0.1179	< 0.0001
TA	0.0010	<0.0001	<0.0001	0.1151	0.4487	<0.0001	<0.0001	0.1031	0.6198	--	0.0133	0.0135	<0.0001	0.0016	0.0047	<0.0001	<0.0001	<0.0001
SW	0.0020	<0.0001	0.0002	0.0001	0.6981	0.0016	<0.0001	0.0947	0.7739	0.0271	--	0.5647	0.0406	0.0852	0.1231	0.0001	<0.0001	<0.0001
ZZ	0.0010	0.0019	0.0002	0.0022	0.2793	0.0052	0.0155	0.0778	0.6750	0.0140	0.6928	--	0.5485	0.4484	0.6091	<0.0001	0.0010	0.0002
HW	<0.0001	<0.0001	<0.0001	<0.0001	0.1517	<0.0001	-	0.0039	0.1436	0.0000	0.0637	0.5391	--	0.6782	0.0528	<0.0001	<0.0001	< 0.0001
CH	<0.0001	<0.0001	<0.0001	<0.0001	0.0254	<0.0001	<0.0001	0.0007	0.0948	0.0008	0.1401	0.3483	0.6043	--	0.2672	<0.0001	<0.0001	<0.0001
SA	<0.0001	<0.0001	<0.0001	<0.0001	0.0977	<0.0001	<0.0001	0.0108	0.1270	0.0033	0.1649	0.6377	0.0481	0.2063	--	<0.0001	<0.0001	< 0.0001
MA	<0.0001	<0.0001	<0.0001	<0.0001	0.2022	<0.0001	<0.0001	<0.0001	0.0659	<0.0001	0.0002	0.0002	<0.0001	<0.0001	<0.0001	--	<0.0001	< 0.0001
KR	<0.0001	<0.0001	<0.0001	<0.0001	0.3893	0.0056	<0.0001	<0.0001	0.1440	<0.0001	<0.0001	0.0024	<0.0001	<0.0001	<0.0001	<0.0001	--	< 0.0001
NA	<0.0001	<0.0001	<0.0001	<0.0001	0.0082	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0002	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	--

	North-central	East	South
North-central	-	<0.0001	<0.0001
East	<0.0001	-	<0.0001
South	<0.0001	<0.0001	-

Genic differentiation can be found above the diagonal and genotypic below. Adjusted alpha-value for 5% nominal level after Bonferroni correction for pairwise comparisons of populations is 0.0003, for pairwise comparisons of regions is 0.0167.

Table 3.8. Population differentiation using FST and RST between each locale and region pair.

	BE	WA	ER	KE	MK	AB	AM	SE	NG	TA	SW	ZZ	HW	CH	SA	MA	KR	NA
BE	-	0.018	0.253	0.016	0.049	0.022	0.039	0.015	0.034	0.040	0.062	0.073	0.070	0.077	0.082	0.053	0.069	0.083
WA	<0.001	-	0.259	0.033	0.061	0.011	0.041	0.044	0.046	0.066	0.071	0.063	0.070	0.072	0.077	0.074	0.060	0.079
ER	<0.001	<0.001	-	0.182	0.217	0.165	0.173	0.203	0.121	0.185	0.160	0.198	0.176	0.178	0.185	0.148	0.158	0.176
KE	0.012	0.032	0.016	-	0.009	0.016	0.021	<0.001	<0.001	0.008	0.012	0.020	0.013	0.022	0.022	0.032	0.020	0.032
MK	0.014	0.013	0.031	0.010	-	<0.001	0.012	0.014	<0.001	0.038	<0.001	0.015	0.018	0.046	0.054	0.036	<0.001	0.041
AB	<0.001	0.009	<0.001	<0.001	<0.001	-	0.009	0.027	0.011	0.039	0.040	0.046	0.045	0.046	0.054	0.039	0.019	0.060
AM	0.120	0.133	0.227	0.031	0.011	0.025	-	0.022	0.011	0.029	0.035	0.027	0.034	0.025	0.026	0.047	0.027	0.060
SE	0.060	0.080	0.151	0.004	0.004	<0.001	<0.001	-	<0.001	0.013	0.011	0.017	0.014	0.022	0.016	0.037	0.034	0.048
NG	0.028	0.054	0.064	<0.001	<0.001	<0.001	<0.001	<0.001	-	0.001	<0.001	<0.001	0.001	0.006	0.009	0.014	0.001	0.019
TA	<0.001	0.036	0.015	<0.001	0.062	<0.001	0.050	0.020	<0.001	-	0.018	0.027	0.018	0.015	0.020	0.029	0.042	0.021
SW	0.076	0.087	0.146	0.011	<0.001	<0.001	<0.001	<0.001	<0.001	0.029	-	-0.004	0.002	0.007	0.007	0.037	0.018	0.020
ZZ	0.244	0.231	0.476	0.126	0.058	0.151	0.056	0.062	0.086	0.187	0.035	-	<0.001	0.002	<0.001	0.046	0.021	0.019
HW	0.112	0.117	0.184	0.023	<0.001	0.020	<0.001	<0.001	<0.001	0.042	<0.001	0.040	-	<0.001	0.003	0.035	0.019	0.019
CH	0.110	0.104	0.202	0.029	0.012	0.028	0.004	<0.001	0.001	0.054	<0.001	0.037	<0.001	-	0.001	0.034	0.028	0.026
SA	0.228	0.237	0.395	0.102	0.042	0.120	0.025	0.034	0.049	0.143	0.021	0.002	0.024	0.037	-	0.046	0.031	0.022
MA	0.062	0.078	0.122	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.010	<0.001	0.081	<0.001	<0.001	0.047	-	0.042	0.048
KR	0.153	0.178	0.266	0.064	<0.001	0.058	0.005	0.010	0.009	0.084	0.004	0.038	0.011	0.027	<0.001	0.015	-	0.039
NA	<0.001	0.018	<0.001	<0.001	0.013	<0.001	0.069	0.033	0.004	<0.001	0.038	0.189	0.055	0.061	0.162	0.020	0.108	-

	North-central	East	South
North-central	-	0.027	0.058
East	0.061	-	0.013
South	0.119	0.008	-

Values for F_{ST} are above and R_{ST} are below the diagonal.

Table 3.9. Analysis of molecular variance (AMOVA).

Source of Variation	df	SS	VC	PV	F-value		P-value
Among Groups	2	23.621	0.028	1.36	F_{CT}	0.014	<0.001
Among Populations Within Groups	15	63.537	0.046	2.25	F_{SC}	0.023	<0.001
Within Populations	910	1794.830	1.972	96.39	F_{ST}	0.036	<0.001
Total	927	1881.988	2.046				

Variables are, degrees of freedom (df), sum of squares (SS), variance component (VC), percent variance (PV), F value (F_{CT}, F_{SC}, F_{ST}) and P-value. Group designations can be found on Table 3.1.

Table 3.10. Calculations to examine the number of African savanna elephant (*Loxodonta africana*) population subdivisions.

Admixture - Correlated

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	3	-11039.63333	0.057735	—	—	—
2	3	-11043.5	0.34641	-3.866667	317.066667	915.292627
3	3	-11364.43333	89.812935	-320.933333	513.733333	5.720037
4	3	-11171.63333	30.900216	192.8	681.233333	22.046232
5	3	-11660.06667	132.646912	-488.433333	256.966667	1.937223
6	3	-11891.53333	73.08039	-231.466667	490.2	6.707682
7	3	-12613.2	265.165741	-721.666667	516	1.945953
8	3	-12818.86667	90.950829	-205.666667	258.5	2.842195
9	3	-13283.03333	496.539881	-464.166667	402.6	0.810811
10	3	-13344.6	203.013005	-61.566667	435.733333	2.146332
11	3	-13841.9	134.301005	-497.3	128.866667	0.959536
12	3	-14210.33333	555.189925	-368.433333	200.3	0.360777
13	3	-14378.46667	498.496403	-168.133333	262.566667	0.526717
14	3	-14809.16667	594.392272	-430.7	1016.166667	1.709589
15	4	-14223.7	256.096089	585.466667	750.733333	2.931452
16	3	-14388.96667	351.675437	-165.266667	141.533333	0.402454
17	3	-14412.7	149.763246	-23.733333	425.633333	2.842041
18	3	-14862.06667	296.768535	-449.366667	—	—

Admixture - Independent

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	3	-11042.16667	0.057735	—	—	—
2	3	-11128.4	3.994997	-86.233333	160.7	40.225313
3	3	-11375.33333	13.91989	-246.933333	160.466667	11.527869
4	3	-11782.73333	28.901442	-407.4	41.866667	1.448601
5	3	-12148.26667	49.590859	-365.533333	33.6	0.677544
6	3	-12480.2	33.846861	-331.933333	165.433333	4.887701
7	3	-12977.56667	38.374514	-497.366667	48.8	1.271677
8	3	-13523.73333	59.861952	-546.166667	43.9	0.733354
9	3	-14026	225.534986	-502.266667	189.433333	0.839929
10	3	-14338.83333	466.777359	-312.833333	87.233333	0.186884
11	3	-14564.43333	89.131046	-225.6	468.666667	5.258175
12	3	-14321.36667	511.282137	243.066667	221.7	0.433616
13	3	-14300	618.140017	21.366667	629.866667	1.018971
14	3	-13648.76667	245.410276	651.233333	167.866667	0.684025
15	3	-12829.66667	121.189493	819.1	600.3	4.9534
16	3	-12610.86667	211.443476	218.8	16.1	0.076143
17	3	-12375.96667	99.165384	234.9	121.133333	1.221528
18	3	-12262.2	47.242883	113.766667	—	—

Table 3.10. Continued.

Noadmixture - Correlated

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	3	-11039.63333	0.057735	—	—	—
2	3	-10990.6	3.318132	49.033333	333.766667	100.588725
3	3	-11275.33333	25.058798	-284.733333	508.066667	20.274982
4	3	-11052	10.935721	223.333333	769.666667	70.38097
5	3	-11598.33333	184.692456	-546.333333	240.133333	1.300179
6	3	-11904.53333	18.097053	-306.2	30.933333	1.709302
7	3	-12241.66667	204.296264	-337.133333	281	1.375453
8	3	-12859.8	563.484907	-618.133333	593.433333	1.053149
9	3	-12884.5	73.430988	-24.7	171.666667	2.337796
10	3	-13080.86667	123.351382	-196.366667	459.266667	3.723239
11	3	-13736.5	814.206209	-655.633333	829.766667	1.019111
12	3	-13562.36667	428.077146	174.133333	233.733333	0.546008
13	3	-13621.96667	103.750775	-59.6	179.366667	1.728822
14	3	-13860.93333	238.063864	-238.966667	298.4	1.253445
15	3	-13801.5	874.726866	59.433333	108.708333	0.124277
16	4	-13850.775	312.597989	-49.275	161.283333	0.515945
17	3	-14061.33333	703.767549	-210.558333	585.791667	0.832365
18	3	-13686.1	225.79081	375.233333	—	—

Noadmixture - Independent

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	3	-11039.63333	0.057735	—	—	—
2	3	-10990.6	3.318132	49.033333	333.766667	100.588725
3	3	-11275.33333	25.058798	-284.733333	508.066667	20.274982
4	3	-11052	10.935721	223.333333	769.666667	70.38097
5	3	-11598.33333	184.692456	-546.333333	240.133333	1.300179
6	3	-11904.53333	18.097053	-306.2	30.933333	1.709302
7	3	-12241.66667	204.296264	-337.133333	281	1.375453
8	3	-12859.8	563.484907	-618.133333	593.433333	1.053149
9	3	-12884.5	73.430988	-24.7	171.666667	2.337796
10	3	-13080.86667	123.351382	-196.366667	459.266667	3.723239
11	3	-13736.5	814.206209	-655.633333	829.766667	1.019111
12	3	-13562.36667	428.077146	174.133333	233.733333	0.546008
13	3	-13621.96667	103.750775	-59.6	179.366667	1.728822
14	3	-13860.93333	238.063864	-238.966667	298.4	1.253445
15	3	-13801.5	874.726866	59.433333	108.708333	0.124277
16	4	-13850.775	312.597989	-49.275	161.283333	0.515945
17	3	-14061.33333	703.767549	-210.558333	585.791667	0.832365
18	3	-13686.1	225.79081	375.233333	—	—

The *ad hoc* method of Evanno et al. (2005) was used to examine the number of population subdivisions for elephants across Africa. The method was implemented in Structure Harvester (Earl & vonHoldt 2012). Calculations utilized the results from STRUCTURE software based on 1 million Markov chain Monte Carlo generations following a burn-in of 100,000 steps and 3 repetitions.

Table 3.11. Calculations to examine the number of African elephant (*Loxodonta*) population subdivisions.

Admixture - Correlated

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	3	-17251.6	0.458258	—	—	—
2	3	-14511.03333	0.763763	2740.566667	2546.9	3334.674868
3	3	-14317.36667	1.357694	193.666667	91.133333	67.123612
4	3	-14214.83333	1.154701	102.533333	43.5	37.672105
5	3	-14155.8	2.306513	59.033333	22.766667	9.870602
6	3	-14119.53333	2.83784	36.266667	35	12.333326
7	3	-14118.26667	6.017752	1.266667	19.4	3.223795
8	3	-14136.4	10.049378	-18.133333	8.766667	0.872359
9	3	-14145.76667	5.53022	-9.366667	0.966667	0.174797
10	3	-14154.16667	9.303942	-8.4	24.833333	2.66912
11	3	-14187.4	21.589118	-33.233333	39.3	1.820362
12	3	-14259.93333	16.519786	-72.533333	3.7	0.223974
13	3	-14328.76667	1.550269	-68.833333	5.533333	3.569274
14	3	-14403.13333	16.163642	-74.366667	7.733333	0.47844
15	3	-14469.76667	17.86962	-66.633333	8.4	0.470072
16	3	-14544.8	29.247051	-75.033333	21.666667	0.740815
17	3	-14641.5	57.161963	-96.7	10.533333	0.184272
18	3	-14727.66667	62.716532	-86.166667	137.433333	2.191341
19	3	-14951.26667	15.267067	-223.6	186.466667	12.213654
20	3	-14988.4	30.831153	-37.133333	67.866667	2.201237
21	3	-15093.4	119.902002	-105	95.8	0.798986
22	3	-15294.2	102.302346	-200.8	62.166667	0.607676
23	3	-15557.16667	71.396242	-262.966667	337.933333	4.733209
24	3	-15482.2	67.000224	74.966667	—	—

The *ad hoc* method of Evanno et al. (2005) was used to examine the number of population subdivisions for elephants across Africa. The method was implemented in Structure Harvester (Earl & vonHoldt 2012). Calculations utilized the results from STRUCTURE software based on 1 million Markov chain Monte Carlo generations following a burn-in of 100,000 steps and 3 repetitions.

Table 3.12. Isolation by distance.

Genetic Distance	Straight Line Distance (km)	r	P
F_{ST}	Genetic distance vs Geographic distance	0.48	< 0.0001
	Genetic distance and log(Geographic distance)	0.43	< 0.0001
	log(Genetic distance) and Geographic distance	0.52	< 0.0001
	log(Genetic distance) and log(Geographic distance)	0.55	< 0.0001
Rousset's distance measure	Genetic distance vs Geographic distance	0.44	< 0.0001
$[F_{ST}/(1-F_{ST})]$	Genetic distance and log(Geographic distance)	0.39	< 0.0001
	log(Genetic distance) and Geographic distance	0.52	< 0.0001
	log(Genetic distance) and log(Geographic distance)	0.56	< 0.0001
Slatkin's Similarity index	Genetic distance vs Geographic distance	-0.05	0.2825
$[M=((1/F_{ST})-1)/4]$	Genetic distance and log(Geographic distance)	-0.06	0.2230
	log(Genetic distance) and Geographic distance	0.10	0.8839
	log(Genetic distance) and log(Geographic distance)	0.15	0.9584

Genetic Distance	Indirect Distance (km)	r	P
F_{ST}	Genetic distance vs Geographic distance	0.45	0.0006
	Genetic distance and log(Geographic distance)	0.43	< 0.0001
	log(Genetic distance) and Geographic distance	0.50	< 0.0001
	log(Genetic distance) and log(Geographic distance)	0.56	< 0.0001
Rousset's distance measure	Genetic distance vs Geographic distance	0.41	0.0037
$[F_{ST}/(1-F_{ST})]$	Genetic distance and log(Geographic distance)	0.40	< 0.0001
	log(Genetic distance) and Geographic distance	0.51	< 0.0001
	log(Genetic distance) and log(Geographic distance)	0.56	< 0.0001
Slatkin's Similarity index	Genetic distance vs Geographic distance	-0.05	0.2558
$[M=((1/F_{ST})-1)/4]$	Genetic distance and log(Geographic distance)	-0.06	0.2084
	log(Genetic distance) and Geographic distance	0.08	0.8266
	log(Genetic distance) and log(Geographic distance)	0.15	0.9542

Straight line geographic distance assuming no barriers to movement. Indirect geographic distance assumes central Africa (tropical forest) as a barrier; therefore, distances were calculated to pass through east Africa (AB population as the waypoint between northern and southern localities).

Table 3.13. Immigration rates of *Loxodonta* africana between each locale and region pair.

	BE	WA	ER	KE	MK	AB	AM	SE	NG	TA	SW	ZZ	HW	CH	SA	MA	KR	NA
BE	**0.68**	0.01	0.01	0.01	0.01	0.01	0.03	0.01	0.01	0.01	0.01	0.01	**0.11**	0.01	0.01	0.01	0.01	0.02
WA	0.01	**0.68**	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	**0.19**	0.01	0.01	0.01	0.01	0.01
ER	0.02	0.02	**0.69**	0.06	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
KE	0.00	0.00	0.00	**0.78**	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	**0.14**	0.00	0.00	0.00	0.00	0.01
MK	0.02	0.02	0.02	0.05	**0.69**	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02
AB	0.01	0.01	0.01	0.03	0.01	**0.68**	0.01	0.01	0.01	0.01	0.01	0.01	**0.15**	0.01	0.01	0.01	0.01	0.01
AM	0.00	0.00	0.00	0.01	0.00	0.00	**0.68**	0.00	0.00	0.00	0.00	0.00	**0.25**	0.00	0.00	0.00	0.00	0.00
SE	0.01	0.01	0.01	0.01	0.01	0.01	0.02	**0.68**	0.01	0.01	0.01	0.01	**0.17**	0.01	0.01	0.01	0.01	0.01
NG	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	**0.68**	0.01	0.01	0.01	**0.16**	0.01	0.01	0.01	0.01	0.01
TA	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	**0.68**	0.01	0.01	**0.19**	0.01	0.01	0.01	0.01	0.01
SW	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	**0.68**	0.01	**0.18**	0.01	0.01	0.01	0.01	0.01
ZZ	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	**0.68**	**0.05**	0.01	0.01	0.01	0.01	**0.11**
HW	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	**0.87**	0.01	0.01	0.01	0.01	0.03
CH	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	**0.21**	**0.67**	0.01	0.01	0.01	0.01
SA	0.01	0.01	0.01	0.06	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	**0.16**	0.01	**0.68**	0.01	0.01	0.01
MA	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	**0.19**	0.01	0.01	**0.68**	0.01	0.01
KR	0.01	0.01	0.01	**0.17**	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.01	0.01	0.01	**0.67**	0.05
NA	0.00	0.01	0.00	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	**0.89**

	North-central	East	South
North-central	**0.69**	**0.27**	0.04
East	0.00	**0.99**	0.01
South	0.00	0.08	**0.92**

Source localities are given in rows, recipient localities in columns. Values in boxes along the diagonal are self-recruitment rates, boldface values are likely immigration sources when the rate is greater than 10%.

Table 3.14. Wilcoxon's sign-ranks test to estimate heterozygote excess and departure from mutation drift equilibrium indicative of a recent bottleneck.

	Mean	Mean	IAM		SMM		TPM	
	N	H_E	1-tailed	2-tailed	1-tailed	2-tailed	1-tailed	2-tailed
BE	17.80*	0.567	**0.006**	**0.012**	0.098	0.195	**0.037**	0.074
WA	42.00	0.516	0.150	0.301	0.820	0.426	0.455	0.910
ER	< 20*	-	-	-	-	-	-	-
KE	116.00	0.547	**0.016**	**0.032**	0.500	1.000	0.188	0.375
MK	< 20*	-	-	-	-	-	-	-
AB	31.80	0.576	**0.001**	**0.002**	0.082	0.164	**0.005**	**0.010**
AM	119.00	0.566	**0.007**	**0.014**	0.754	0.557	0.138	0.275
SE	35.60	0.549	0.246	0.492	0.539	1.000	0.313	0.625
NG	28.00	0.568	0.161	0.322	0.385	0.770	0.188	0.375
TA	43.40	0.551	**0.012**	**0.024**	0.423	0.846	0.116	0.232
SW	35.40	0.545	0.116	0.232	0.313	0.625	0.246	0.492
ZZ	22.20	0.520	0.116	0.232	0.385	0.770	0.246	0.492
HW	76.40	0.535	0.065	0.131	0.500	1.000	0.116	0.232
CH	56.40	0.537	0.065	0.131	0.138	0.275	0.097	0.193
SA	48.00	0.522	0.161	0.322	0.813	0.432	0.461	0.922
MA	42.00	0.605	**0.001**	**0.003**	0.161	0.322	0.042	0.084
KR	79.40	0.533	0.065	0.131	0.423	0.846	0.188	0.375
NA	95.00	0.544	**0.001**	**0.002**	0.577	0.922	0.053	0.105
ALL	898.40	0.563	0.080	0.160	0.999	**0.005**	0.688	0.695

Variables are (N) number of gene copies, (H_E) heterozygosity, (IAM) infinite allele model, (SMM) step-wise mutation model, and (TPM) two phase model. P-values are for Wilcoxon's sign-ranks (1-tailed) for heterozygote excess and (2-tailed) departures from mutation-drift equilibrium. Values in bold are significant ($P < 0.05$). *Average number of gene copies was lower than acceptable (minimum of 20) for this analysis. ER and MK were excluded due to extremely low number of gene copies (average < 3); however individuals in these localities were included for analysis across all localities.

CHAPTER 4: THE ELEPHANTS OF GASH-BARKA, ERITREA: NUCLEAR AND MITOCHONDRIAL GENETIC PATTERNS

Abstract

Eritrea marks the northern range boundary for African elephants. Only about 100 elephants persist in the Gash-Barka administrative zone. Elephants in Eritrea have become completely isolated, with no gene flow from other elephant populations. The conservation of Eritrean elephants would benefit from an understanding of their genetic affinities to elephants elsewhere on the continent, and of the degree to which genetic variation persists in the population. Using dung samples from Eritrean elephants, we examined 18 species-diagnostic single nucleotide polymorphisms in 3 nuclear genes, sequences of mitochondrial *HVR1* and *ND5*, and genotyped 11 microsatellite loci. All sampled Eritrean elephants carried nuclear and mtDNA markers establishing them as savanna elephants, with closer genetic affinity to Eastern than to North-Central savanna elephant populations, and contrary to speculation by some scholars that forest elephants were found in Eritrea. Mitochondrial DNA diversity was relatively low, while two haplotypes unique to Eritrea predominated. STR genotypes could only be determined for a small number of elephants, but suggested that the population suffers from low diversity. Conservation efforts should aim to protect Eritrean elephants and their habitat in the short run, with restoration of habitat connectivity and genetic diversity as long-term goals.

Introduction

Eritrea marks the northern boundary of current African elephant distribution, with elephants persisting in a small fragment of their formerly extensive range. African elephants were once found throughout Eritrea, but by the early 20th century they were believed to have been extirpated (Gowers 1948), although a small population was found to have persisted at low population densities in the southwest (Hagos *et al.* 2003; Yalden *et al.* 1986). The current distribution of elephants is restricted to a 5,293 km^2 area of land in the Zoba Gash-Barka, one of the six administrative zones, located in the southern part of western Eritrea (Figure 4.1) (Blanc *et al.* 2007). Gash-Barka is a dry region with habitat consisting mostly of doum palm, ziziphus bush, acacia woodland and open grassland savanna (Hagos *et al.* 2003). Most surveys report sightings of only a few individuals (Barnes *et al.* 1999; Litoroh 1997). However, one estimate suggested that 100 to 200 elephants persisted in the 1950s (Largen & Yalden 1987; Leuenberger 1955), while a recent aerial survey conducted in Gash-Barka between 2001 and 2003 estimated that ca. 100 African elephants remain in Eritrea (Hagos *et al.* 2003; Shoshani *et al.* 2004). In 2012, the government of Eritrea indicated that the numbers and range of elephants appear to be increasing, and that ca. 120 elephants persist (Anonymous 2013). Protecting elephant habitat is considered by the government to be a priority for biodiversity conservation (Weldeyohannes & Siratu 2010).

Within Eritrea, the geographic range of elephants is approximately 4,200 km^2 which includes narrow corridors connecting the northern and southern extents of their range (Yacob *et al.* 2004). During the wet season, some Eritrean elephants migrate into northern Ethiopia (Shoshani *et al.* 2000; Shoshani *et al.* 2004), utilizing the additional range within the Tkezze Valley Wildlife Reserve, which is 1,130 km^2 (Blanc *et al.* 2007). The study of these elephants is made difficult by their migrations between Eritrea and Ethiopia. During a 27 month study, the mortality rate was estimated to be 4.9% per year, which is comparable to the 14 year average of 4.71% in the Samburu elephant population in northern Kenya (Wittemyer *et al.* 2013) and less than the mortality rate (17.1 % of juveniles or 10.5% of adults) that has been estimated as necessary to prevent population growth in savanna elephants (Woolley *et al.* 2008). Many elephant deaths in Eritrea are attributable to the human presence in the area, although ivory poaching has not been of major concern (Yacob *et al.* 2004). The Eritrean groups observed

included a substantial proportion of infant and sub-adult individuals (Hagos *et al.* 2003; Shoshani *et al.* 2004).

Elephants in the region are believed to be isolated, with the nearest other elephant population over 400 km away (Blanc *et al.* 2007). Eritrean elephants are thus vulnerable to a decrease in fitness due to inbreeding and loss of genetic variation (Reed & Frankham 2003). Understanding the genetic diversity and affinities of this population, and determining the effects of limited gene flow, can contribute to scientifically sound conservation practices to ensure their long-term persistence. We therefore examined genetic markers previously characterized in elephants across Africa to examine the genetic diversity and affinities of Eritrean elephants.

Material and Methods

Samples

This study was conducted in compliance with the University of Illinois Institutional Animal Care and Use Committed (IACUC) approved protocol number 09036. Samples were obtained in full compliance with required permits. Thirty-three dung samples were collected from elephants in the Gash-Barka region of Eritrea between 2001 and 2003 and stored in "blue" alcohol (ethanol with methanol additive). Dung was collected as part of the elephant population census reported by Shoshani and others (2004). The dung was collected in various geographic locations during this time period (Table 4.1, Figure 4.2). To minimize the possibility of duplicate collection of samples, exclusion criteria were used consisting of (1) ability of elephants to travel distances given differences in time of sample collection; (2) similarity of herd or group composition; and (3) similarity of markings including ear and body marks, tusk characters, and soleprints (Hagos *et al.* 2003; Shoshani *et al.* 2004). A total of 83 distinct elephants were counted during the census, though an additional 45 sightings of elephants were excluded from the census total using these conservative exclusion criteria. None of the samples that were successfully amplified (below) were from the potential duplicates, thus all sequences and genotypes were from distinct individuals.

Mitochondrial and Nuclear DNA Amplification and Sequencing

DNA was extracted using the QIAamp DNA Stool Kit (Qiagen Inc., Valencia, CA) following the recommended protocol. Several DNA markers were unable to be amplified despite repeated attempts and utilization of techniques that typically increase PCR success rate; thus limiting the analyses possible for some individuals. Two regions of the mitochondrial genome were amplified and sequenced. A 319 bp region of the mitochondrial *NADH dehydrogenase 5* (*ND5*) was amplified as previously described (Roca *et al.* 2005). A 314 bp region of the *hyper variable region 1* (*HVR1*) was amplified in two overlapping segments using a combination of four primers developed for low quality DNA, CR-F1 (TGGTCTTGTAAGCCATAAATGAAA) with CR-R1 (GCTTTAATGTGCTATGTAAGACTATG), and CR-F2 (TCGTGCATCACATTATTTACCC) with CR-R2 (TGGTCCTGAAGAAAGAACCAG). PCR was run with an initial step of 95°C for 9:45 min; with cycles of 20 sec at 94°C; followed by 30 sec at 60°C (first 3 cycles), 58°C (next 5 cycles), 56°C (5 cycles), 54°C (5 cycles), 52°C (5 cycles), or 50°C (final 22 cycles); followed by 30 second extension at 72°C; with a final extension after the last cycle of 7 min at 72°C. Short species-diagnostic regions of nuclear DNA sequences for genes *Biglycan* (*BGN*), *Phosphorylase kinase alpha subunit 2* (*PHKA2*) and *Proteolipid protein 1 (PLP*) were amplified following methods previously described (Ishida *et al.* 2011a). All products were enzyme-purified (Hanke & Wink 1994) then sequenced using the BigDye Terminator system (ABI), purified, and resolved at the University of Illinois at Urbana-Champaign Core Sequencing Facility. The software *Sequencher* (Gene Codes Corporation, Ann Arbor, MI) was used to edit and concatenate sequences. There were no indications of nuclear DNA sequences of mitochondrial origin (numts) among the results (Brandt *et al.* 2012; Roca *et al.* 2007). Sequences of four novel mtDNA haplotypes were submitted to GenBank (KC608163-KC608166).

Haplotype Analyses

Mitochondrial DNA sequences were aligned using CLUSTALW 2.0 (Larkin *et al.* 2007) with default parameters, in EBI Web Services (McWilliam *et al.* 2009); alignment output was visually inspected. Haplotype diversity indices were calculated with ARLEQUIN v.3.5 (Excoffier & Lischer 2010). *HVR1* sequences were combined with a larger dataset (Ishida *et al.*

2013) and weighted maximum likelihood distances were used to generate a median joining network using the software NETWORK v.4.6.1 (Bandelt *et al.* 1999).

SNP Analyses

The identities of nuclear DNA sequences were established using NCBI BLAST (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) and were compared to published DNA sequences from savanna elephant, forest elephant and Asian elephant (Roca *et al.* 2005). Species-specific diagnostic nucleotide sites for each gene (*BGN*, *PHK* and *PLP*) were examined, as described previously (Ishida *et al.* 2011a). Since the sex of the individuals was not known, nuclear amplicons were conservatively estimated as representative of 1 rather than 2 X-chromosomes.

Microsatellites

Eight microsatellite loci developed in savanna elephants: LAF10, LAF11, LAF12, LAF13, LAF29, LAF37, LaT05, and LaT06 (Archie *et al.* 2003; Ishida *et al.* 2011b) and 3 loci developed in Asian elephants: EMX3, EMX4, and EMX5 (Fernando *et al.* 2001), were amplified in the Eritrean samples by PCR. Primers were tagged for fluorescence detection (Boutin-Ganache *et al.* 2001) and amplification followed a touchdown thermocycle profile previously described (Ishida *et al.* 2011b; Menotti-Raymond *et al.* 2005). Samples were genotyped on an ABI 3100 Genetic Analyzer and scored using GeneScan 3.7 and Genotyper 2.5 software (Applied Biosystems); alleles were subsequently binned using Allelogram (Morin *et al.* 2009). To verify genotypes and to check for allelic dropout or false alleles, samples were genotyped at least 3 additional times; in no cases were allelic dropouts detected. PCR mixes included 10-25 ug bovine serum albumin. Positive and negative controls were run. In fulfillment of data archiving guidelines (Baker 2013), we have deposited the primary data underlying these analyses with Dryad.

The results for Eritrea are reported here for the first time. One amplification for the Eritrean elephants was generated concurrent with and alongside an additional 555 African elephants from 23 localities that included both savanna elephants (*Loxodonta africana*) and forest elephants (*L. cyclotis*), and followed procedures described therein (Ishida *et al.* 2011b), thus ensuring consistency in platforms and in allelic size comparisons. Diversity indices for microsatellites were calculated using ARLEQUIN v.3.5 (Excoffier & Lischer 2010) and

population structure was examined using the software STRUCTURE 2.3.3 (Hubisz *et al.* 2009). Four models (Pritchard *et al.* 2000) were used to examine the effects of various combinations of assumptions of individual genetic ancestry and genetic relatedness among populations: 1) admixture with correlated allele frequencies, 2) admixture with independent allele frequencies, 3) no admixture with correlated allele frequencies and 4) no admixture with independent allele frequencies. Each model was run 3 times using values of K (possible number of populations) between 1 and 24 genetic clusters, which is the maximum number of putative populations assigned *a priori*. Each analysis was run for a minimum of 1 million Markov chain Monte Carlo steps following a burn-in of at least 100,000 steps. The uppermost hierarchical level of population structure was examined using the *ad hoc* statistic delta K based on the rate of change in ln P (D) between successive K values (Evanno *et al.* 2005), implemented in Structure Harvester (Earl & vonHoldt 2012).

To identify the genetic affinity of Eritrean elephants, STR data from savanna elephants from north central and east Africa were combined with data from Eritrean elephants, and parameters of the STRUCTURE software were modified to allow for "learning samples." Default parameters were used for migrant priors, allele frequencies were only updated from north central and eastern savanna elephant populations, and an admixture model with correlated allele frequencies and K=2 clusters was assumed.

Results

A total of 33 elephant dung samples were collected in the Gash-Barka zone of Eritrea. DNA extraction was attempted at least twice on all samples. The DNA proved to be of low quality, since for 10 samples amplification was never successful for any locus. Sequences of mtDNA were obtained for 21 samples and short nuclear fragments were sequenced for 9 samples. However, only 3 individuals successfully amplified for at least seven of the STR loci. Results were confirmed by repeated genotyping or sequencing.

Mitochondrial Haplotypes

We successfully sequenced mtDNA *ND5* in 20 samples. We identified a single *ND5* haplotype (GenBank accession number KC608166) for all; an NCBI BLAST search revealed that this haplotype occurred in savanna elephants throughout eastern and southern Africa (Ishida *et al.* 2013; Roca *et al.* 2005). Among 15 Eritrean elephant samples that were successfully amplified and sequenced for *HVR1*, three unique haplotypes were detected. These haplotypes were distinguished by only 2 polymorphic nucleotide sites, so that the Eritrean elephant *HVR1* haplotypes differed by 1 nucleotide character state each (Figure 4.3). Thus at *HVR1* the Eritrean elephants had low nucleotide diversity ($\pi = 0.0015$; S.D. +/- 0.0016), and low haplotype diversity ($h = 0.4476$; S.D. +/- 0.1345). The three unique *HVR1* Eritrean haplotypes were aligned with those previously published for African elephants, and used to generate a median joining network (Figure 4.3A). Eight mtDNA groups had previously been identified by Ishida and colleagues (Ishida *et al.* 2013); the 3 Eritrean haplotypes were all within the "savanna wide" group (Figure 2). One haplotype (found in 3 Eritrean samples) was identical to a previously reported haplotype (Figure 4.3; GenBank accession number AY741325) (Nyakaana *et al.* 2002), which occurs in elephants from across Eastern and Southern Africa (Debruyne 2005; Eggert *et al.* 2002; Johnson *et al.* 2007; Nyakaana *et al.* 2002). The remaining 2 haplotypes (found in 1 and 11 samples) were novel and confined to Eritrean elephants; they did not match sequences reported by any previous study (Figure 4.3).

Nuclear SNPs

Three X-linked nuclear genes (*BGN*, *PHK*, *PLP*) have nucleotide sites with fixed character states that distinguish forest from savanna elephants (Roca *et al.* 2005). Primers for PCR amplification of very short fragments containing one or more of these diagnostic sites have previously been developed for use with degraded DNA (Ishida *et al.* 2011a). Using these primers, we obtained at least one of the genic sequences for 9 of the Eritrean samples (Table 4.2), generating a total of 21 unlinked sequences with sites that distinguish forest from savanna elephants. At every one of these segments, savanna elephant-specific nucleotide character states were present (Table 4.2); sequences with sites that matched a character state typical of forest elephants were never found (Table 4.2). Fisher's exact tests comparing these sites established that there were no significant differences ($p \approx 1.00$) in the proportions of character states between

Eritrean elephants and savanna elephants (Ishida *et al.* 2011a) (Table 4.3). By contrast, a Fisher's exact test found highly significant differences ($p < 10^{-4}$) between the character states found in Eritrean elephants and those in previously examined forest elephants (Ishida *et al.* 2011a).

Microsatellites

For microsatellites, only 3 elephants from Eritrea were successfully genotyped: 2 at 8 loci, and 1 at 7 loci. The low success rate may be attributable to degraded DNA, perhaps due to field or storage conditions. Allele scores were confirmed by repeated PCR and genotyping (at least 3 replicates), no allelic dropouts were detected. Within Eritrea, 5 of the loci were polymorphic and three were monomorphic, with an average number of alleles per locus of 1.46. Observed and expected heterozygosity were 0.36 and 0.29, respectively; F_{IS} was -0.37 and all polymorphic loci were in Hardy-Weinberg equilibrium. For these same loci, among savanna elephants from across the African continent genotyped by a previous study (Ishida *et al.* 2011b), the average number of alleles per locus had been 9.88, while observed and expected heterozygosity had been 0.57 and 0.58 respectively. To account for the small sample size from Eritrea, 3 individuals from each non-Eritrean African savanna elephant population were randomly chosen for analysis. In this analysis, elephants from the rest of Africa were still more diverse than Eritrean elephants: the average number of alleles per locus per population (for sample size n=3) was 1.97 (standard deviation of 0.21) while observed and expected heterozygosity were 0.54 and 0.59, respectively.

Bayesian clustering analysis was performed using STRUCTURE (Pritchard *et al.* 2000) for two data sets that combined the Eritrean individuals with genotypes from a larger group of elephants that had been previously reported (Ishida *et al.* 2011b). The analysis included 555 forest and savanna elephants from outside Eritrea. This supported splitting Africa's elephants into two clusters (K = 2; Figure 4.4A, Figure 4.5 and Table 4.4), one corresponding to African forest elephants, the other to African savanna elephants (Ishida *et al.* 2011b). Partitioning of the 3 Eritrean elephants identified them as savanna elephants (Figure 4.4B). The overall proportion of the Eritrean elephants assigned to the forest elephant partition was 0.08. This partitioning appeared to reflect local differences in savanna elephant allele frequencies, rather than admixture from forest elephants. We examined the data closely, finding three alleles present in Eritrea that were more common in forest than savanna elephants (one allele at each of the loci LAF37,

LaT06 and EMX4). These alleles occurred at high frequencies or were fixed in Eritrean elephants. Even so, these three alleles were also present in other savanna elephants, and no allele at any locus in Eritrean elephants fell outside the size range expected of savanna elephants. No allele in Eritrea had a size that was typical of only forest and not savanna elephants, in cases where the allelic size ranges vary between the species (Ishida *et al.* 2011b). With the caveat that DNA from only three individuals amplified, the close examination of STR allele sizes in Eritrea failed to find evidence for this population having any alleles that would be indicative of admixture from forest elephants.

A second STRUCTURE analysis included the three Eritrean elephants along with previously published genotypes of only savanna elephants from north central Africa (Cameroon) and from east Africa, in order to examine whether Eritrean elephants genetically had a closer affinity to elephants from one region or the other. The parameters of the STRUCTURE software were modified to allow for "learning samples" in which north central and eastern savanna elephants were *a priori* assigned to their known region of origin. The Eritrean elephants were not defined *a priori* as belonging to a population or region. Despite this modification, partitioning between elephants in the two regions was not complete, presumably due to limited differentiation between north central and eastern savanna elephants. Different patterns between the north central and the eastern savanna elephants were evident (Figure 4.4C). The patterns observed among Eritrean elephants more closely resembled those of eastern than those of north-central savanna elephants (Figure 3C). This suggests that Eritrean elephants have a greater nuclear genetic affinity with East African than with North-Central African savanna elephants, consistent with the finding that some Eritrean elephants share control region haplotypes with populations in Eastern Africa, but are not known to share mtDNA haplotypes with Cameroon elephants (Figure 4.3).

Discussion

Some scholars have speculated that war elephants used in the 3rd century BCE that had been captured in Eritrea were forest rather than savanna elephants (Gowers 1948). This was based on a written account of the battle of Raphia in 217 BCE, fought between the armies of Ptolemy IV and Antiochus III during the Syrian Wars, and in which African and Asian elephants

met in combat. The Asian elephants used by Antiochus are described as superior in size and strength over Ptolemy's African elephants (Polybius 1923). Since African savanna elephants are larger than Asian elephants, some writers were led to conclude that the elephants used by Ptolemy could have been African forest elephants (Gowers 1948), which are smaller and have a more compact build than savanna elephants (Grubb *et al.* 2000).

Sequencing of diagnostic single nucleotide polymorphisms found only savanna elephant and not forest elephant nuclear genotypes among the elephants of Eritrea (Figure 4.4, Table 4.2). Also, in Eritrea we detected only S clade mitochondrial DNA, which is carried only by savanna elephants (Ishida *et al.* 2011b). The mtDNA results may be especially telling, because savanna elephants in eastern, southern, and north-central Africa often carry F clade mitochondrial haplotypes that are geographically persistent and may record the ancient presence of forest elephants in a locality (Roca *et al.* 2005). The forest elephant mtDNA may be geographically persistent since female African elephants are non-dispersing (Ishida *et al.* 2011b). Although our results cannot completely rule out the possibility that forest elephants may have existed somewhere in Eritrea in the past, our data provide no support for this speculation. Eritrean elephants comprise a savanna elephant population in which even the forest-derived F clade mtDNA carried by many other savanna elephant populations was not detected. While not consistent with previous speculation about the taxonomic affinity of Eritrean war elephants, our results are consistent with the view that two and only two species of elephant occur in Africa, and that currently hybrids between them are confined to a relatively narrow contact zone between forest and savanna habitats (Ishida *et al.* 2011b). Likewise, our results should dispense with rumors that Asian elephants brought to Eritrea in 1868 had admixed with African elephants in the region (Hagos *et al.* 2003; Shoshani *et al.* 2004).

Both nuclear and mitochondrial data support a closer relationship of Eritrean elephants to savanna elephants in East Africa than to savanna elephants of the north-central Sudanian/Sahelian region (Figures 5.3, 5.4). Of three *HVR1* mitochondrial DNA haplotypes carried by elephants in Eritrea, one is widespread, occurring throughout eastern and southern Africa. The remaining 2 haplotypes are restricted to Eritrea, but differ by a single nucleotide from haplotypes found in eastern but not north-central Africa (Figure 4.3). The single *ND5* haplotype present in Eritrea has also been detected across eastern and southern Africa but not north central Africa (Ishida *et al.* 2013; Roca *et al.* 2005). Although nuclear microsatellite

genotypes were only successful for three Eritrean elephants, each of these had a closer genetic affinity to eastern savanna elephants than to the elephants of Cameroon (Figure 4.4).

Mitochondrial haplotype and nucleotide diversity were both low compared to other savanna elephant populations. For elephants across Africa, average *HVR1* haplotype diversity has been reported as 0.985 (Johnson *et al.* 2007) or 0.85 (Nyakaana *et al.* 2002), about twice the 0.45 of Eritrea. Mitochondrial nucleotide diversity has been reported as 0.030 (Johnson *et al.* 2007) or 0.02 (Nyakaana *et al.* 2002), also about twice the 0.0015 observed in Eritrea. Variation among haplotypes in Eritrea was low, as the three unique *HVR1* haplotypes were defined by two nucleotide variable sites, and only a single *ND5* haplotype was detected. Nuclear diversity was also very low, a previous study reported observed heterozygosity among all African elephant for the same STR loci as 0.50 (Ishida *et al.* 2011b) , higher than the 0.36 observed in Eritrea.
The population of elephants in Eritrea is small; human-wildlife conflicts and habitat loss are major concerns. Currently elephant migration into Ethiopia occurs only during the wet season (Yacob *et al.* 2004). This emphasizes the importance of the habitat in Eritrea for sustaining this population. An increase in suitable and protected habitat may be helpful to the long-term survival of Eritrean elephants. The Agriculture Ministry of Eritrea is committed to preservation of the elephants while minimizing human-elephant conflict, and recently reported that the numbers and range of elephants appear to be increasing (Anonymous 2013). Since the elephant population of Eritrea is small and isolated, it may in the future require genetic management. In the absence of habitat corridors that enable gene flow, genetic management or restoration may eventually become necessary, in which case our results suggest that the Eritrean population would best be augmented using individuals from eastern Africa.

Figures and Tables

Figure 4.1. Map of Eritrea showing current African elephant distribution.

The shaded region pointed to by the arrow indicates the current range of elephants in Eritrea (Litoroh 1997; Shoshani *et al.* 2004). This map is from the IUCN African Elephant Specialist Group – African Elephant Status Report (Blanc *et al.* 2007), which permits reproduction for educational purposes.

Figure 4.1. Continued.

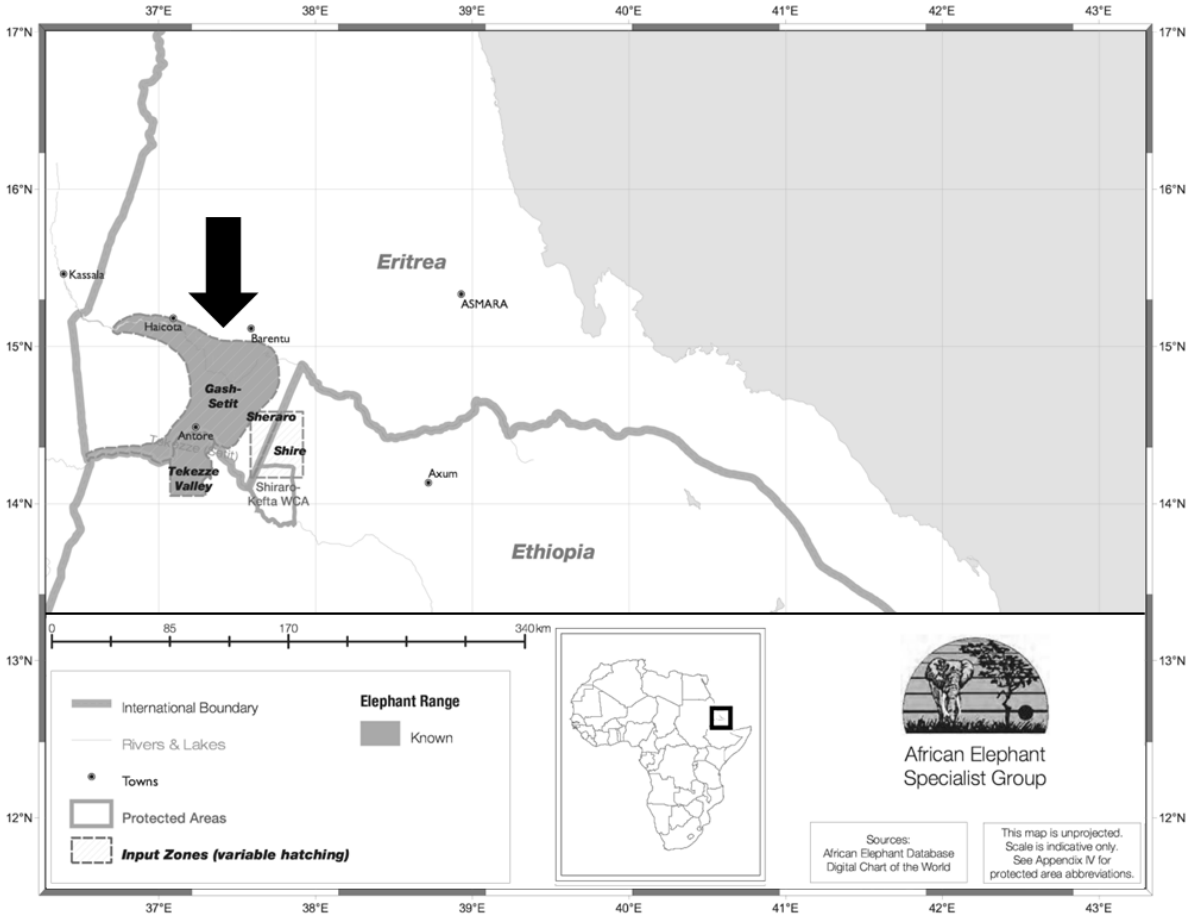


Figure 4.2. Map of Eritrea showing sampling locations.

This map of Gash-Barka, Eritrea shows the locations where dung samples were collected, which were approximated by the name of the nearest town or village (dots) (Shoshani *et al.* 2004). Further details on sample collection can be found in Table 4.1.

Figure 4.2. Continued.

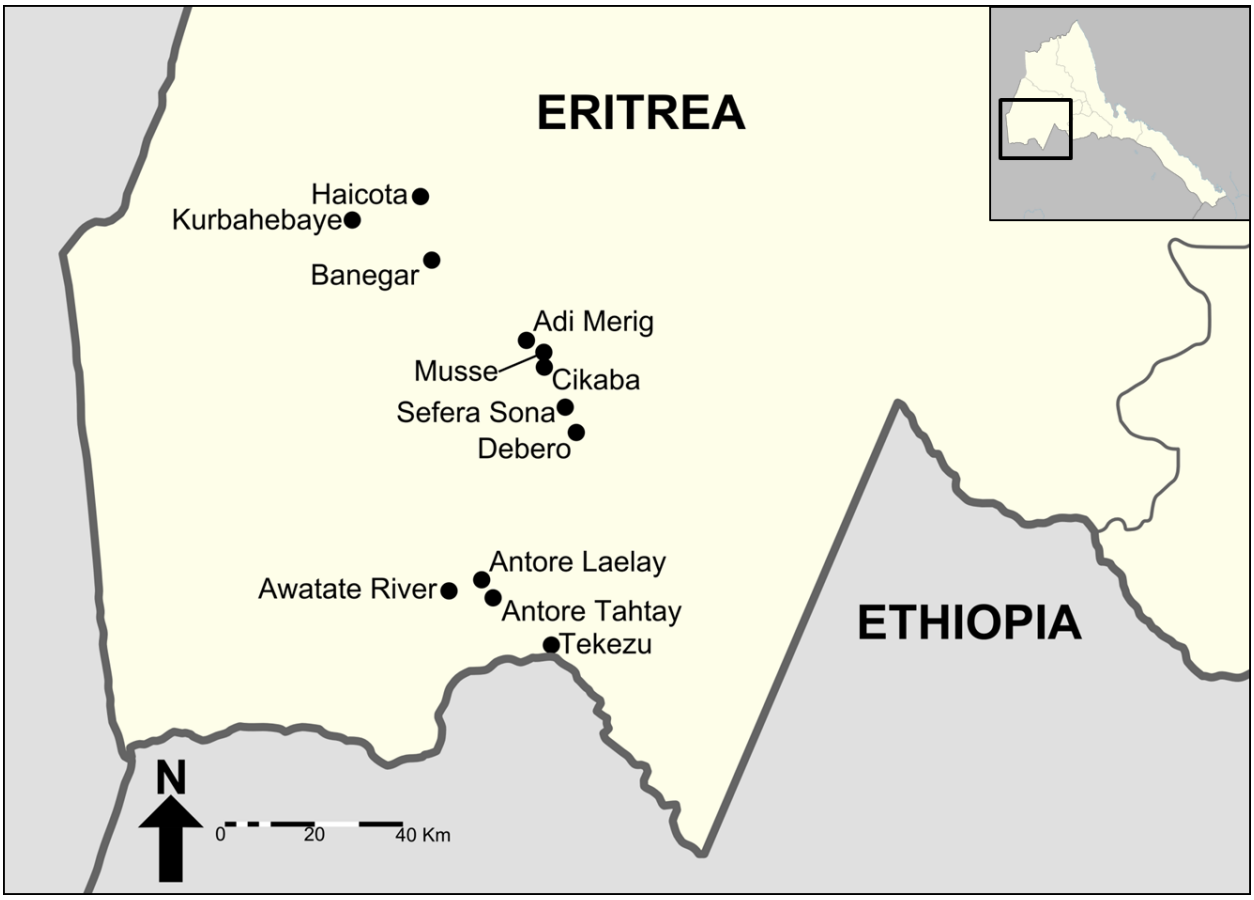


Figure 4.3. Median-joining network of 314 bp mitochondrial HVR1 haplotypes in Eritrean and other African elephants.

(A) The eight mitochondrial groups reported by Ishida et al. (2013) are color-coded. Haplotypes in the groups within the box are carried only by savanna and not forest elephants. (B) Close-up of three groups, showing the haplotypes carried by Eritrean elephants (numbered 1, 2 and 3), for which circle size is proportional to the number of individuals from Eritrea carrying each haplotype. Haplotypes 1 and 2 (Genbank accessions KC608163 and KC608165, respectively) were detected only in Eritrea; haplotype 3 was found in Eritrea and elsewhere. Haplotypes 4-9 differ by a single nucleotide from those present in Eritrea. Haplotypes 3 through 9 have Genbank accessions AY741325, AY742801, AF106236, AF106226, AF106239, AY741074, and AF106235 (Debruyne 2005; Nyakaana *et al.* 2002).

Figure 4.3. Continued.

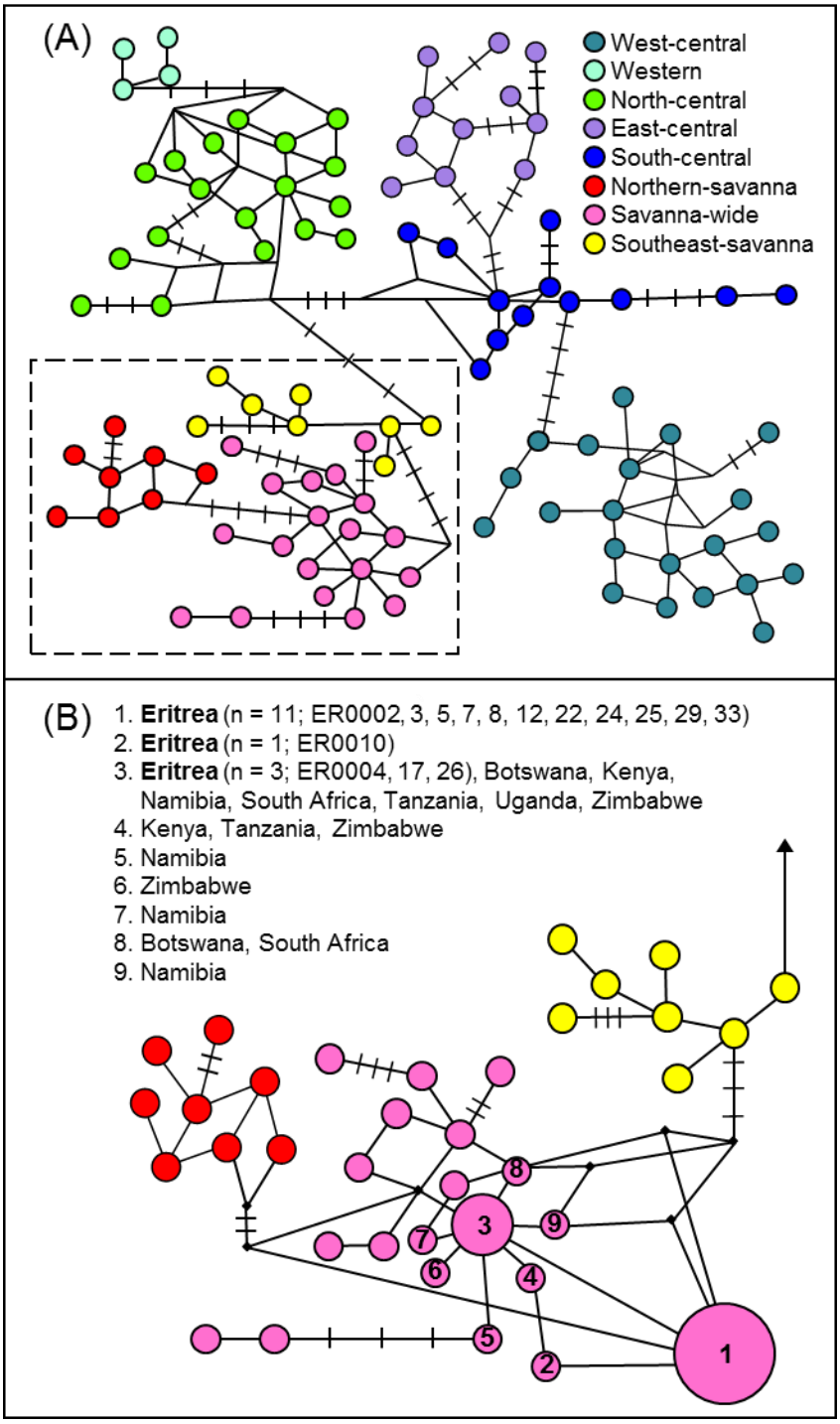


Figure 4.4. Genetic subdivisions among African elephants.

Multi-locus genotype data were used to estimate population subdivision using the program STRUCTURE (Pritchard *et al.* 2000). (A) Two partitions that distinguish African forest elephants (white partition) from African savanna elephants (shaded partition) were successfully reconstructed (Ishida *et al.* 2011b). Three Eritrean (ER) elephants were assigned largely but not completely to the savanna elephant partition. (B) A close-up of STRUCTURE partitioning for the Eritrean elephants. Average assignment of Eritrean elephants to the partition corresponding to forest elephants was ca. 0.08, although this appears to result from drift and not from any admixture by forest elephants. (C) The genetic affinity of Eritrean elephants to other African savanna elephants was examined in STRUCTURE, with non-Eritrean individuals assigned to pre-defined populations as "learning samples". Savanna elephant from Cameroon (BE, WA) were defined as belonging to a north central population while savanna elephants from Kenya and Tanzania (KE, MK, AB, AM, SE, NG, TA) were defined as belonging to an eastern population. However, the two partitions (lightly and darkly shaded) did not completely distinguish geographic regions (Cameroon and East Africa). The Eritrean elephants displayed patterns more closely resembling those of East African than of Cameroon elephants, indicating that Eritrean elephants were genetically more similar to the former. Abbreviations for elephant localities are listed in Materials and Methods.

Figure 4.4. Continued.

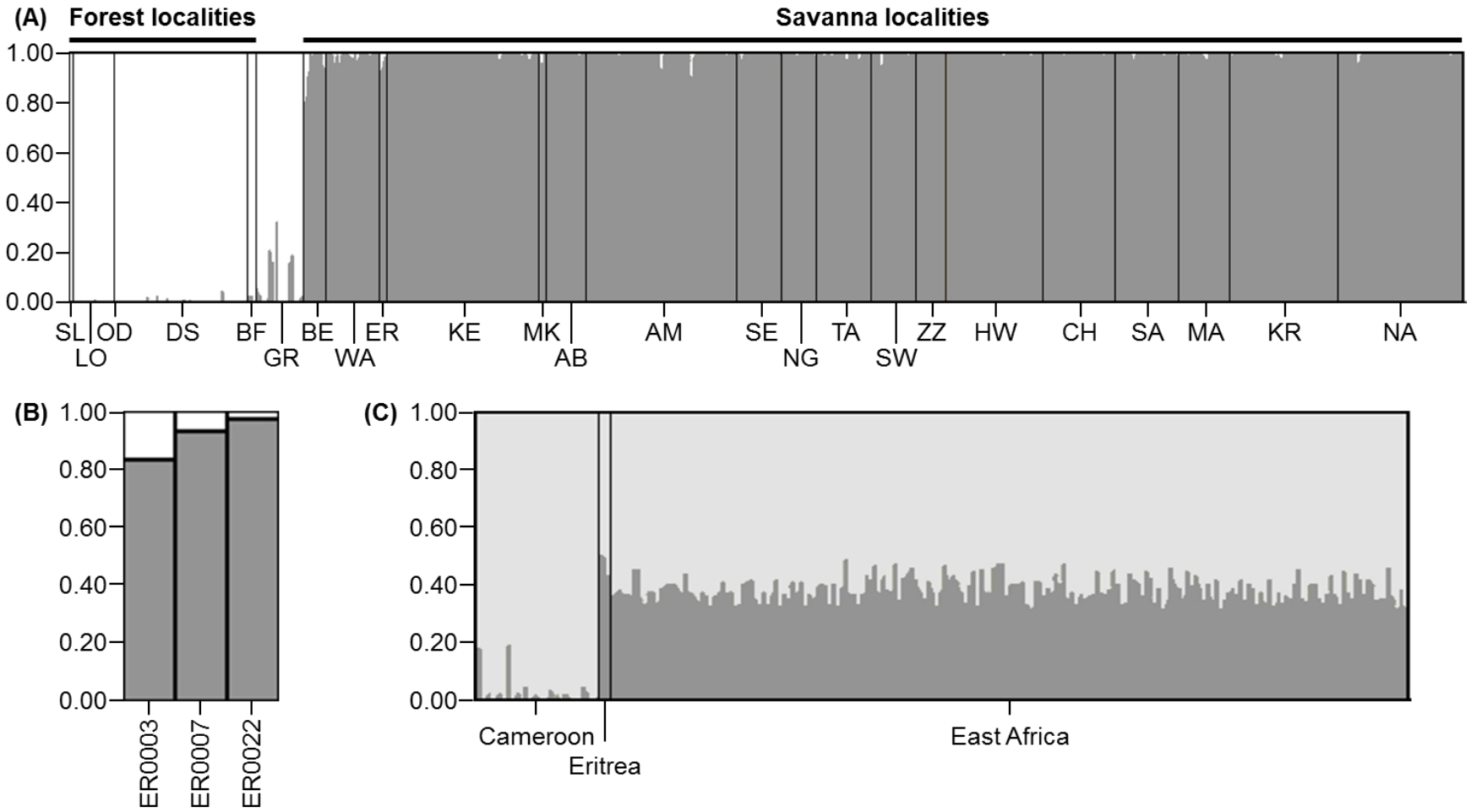


Figure 4.5. Graphical plots for examining the number of genetic subdivisions (K) in African elephants (genus *Loxodonta*).

Results shown are based on the *ad hoc* method of Evanno (Evanno *et al.* 2005) as implemented in Structure Harvester (Earl & vonHoldt 2012). Combinations of assumptions of individual genetic ancestry and genetic relatedness among populations were tested: A) admixture with correlated allele frequencies, B) admixture with independent allele frequencies, C) no admixture with correlated allele frequencies and D) no admixture with independent allele frequencies. For the approach assuming no admixture and independent allele frequencies, the standard deviation of Ln(K) was less than 10^{-7} and therefore K could not be estimated. For all other approaches, the Evanno method supported K = 2 clusters.

Figure 4.5. Continued.

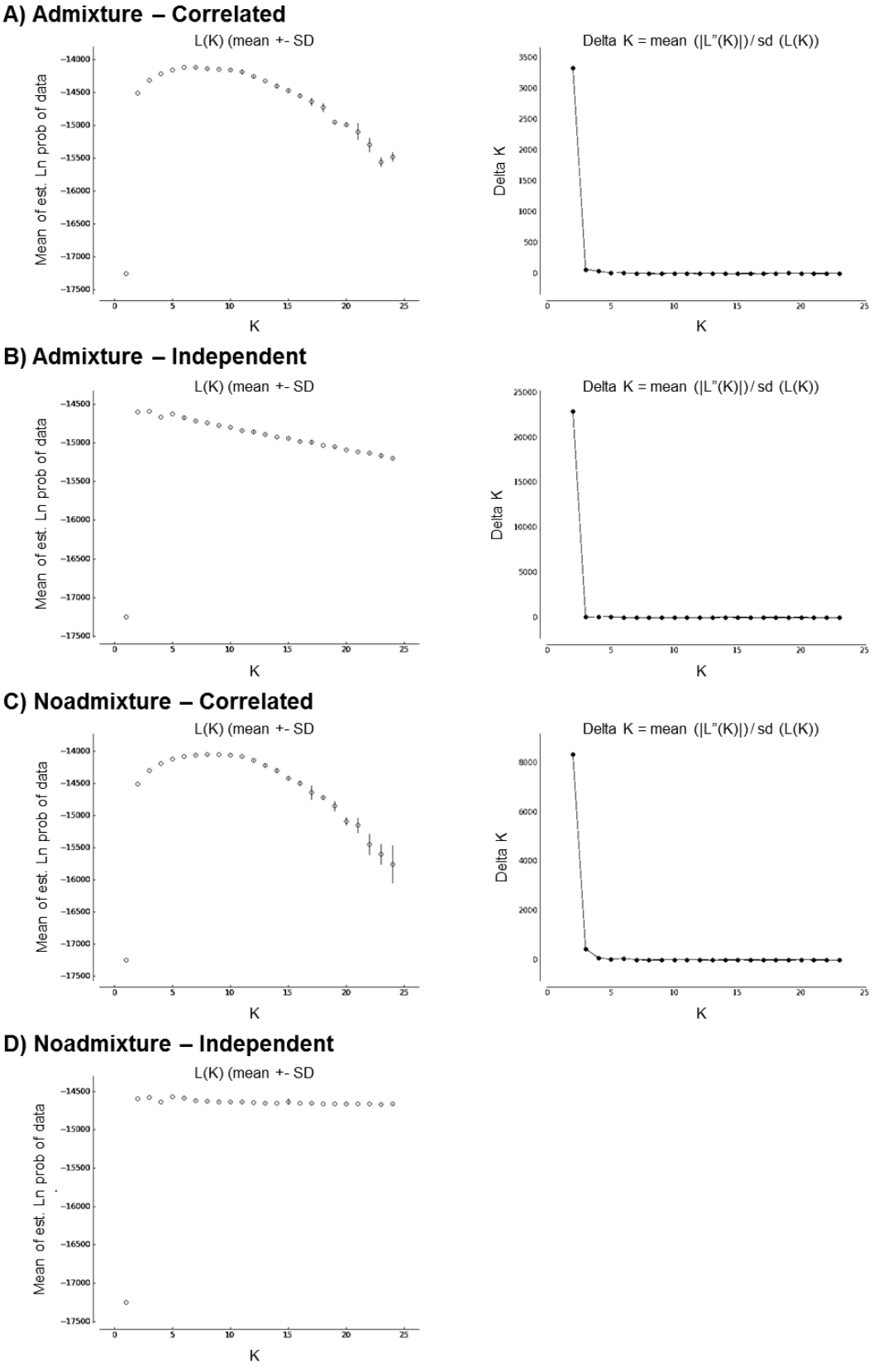


Table 4.1. Sampling data for Eritrean elephants.

Sample ID	*Date Collected*	*Nearest Town / Village*
ER0001	2-Feb-02	Tekezu
ER0002	24-Dec-01	Banegar
ER0003	24-Dec-01	Banegar
ER0004	21-Apr-01	Debero
ER0005	21-Apr-01	Antore Laelay
ER0006	23-Jun-03	Tekezu
ER0007	22-Nov-02	Awtate River
ER0008	27-Dec-03	Haricota
ER0009	Unknown	Unknown
ER0010	2-Feb-03	Kurkahebaye
ER0011	9-May-03	Sefera Sona
ER0012	27-Jan-03	Adi Merig
ER0017	10-Feb-03	Antore Tahtay
ER0022	7-May-03	Musse
ER0024	12-Apr-03	Sefera Sona
ER0025	8-Nov-02	Tekezu
ER0026	22-Nov-02	Awtate River
ER0027	14-Mar-02	Debero
ER0029	13-Jun-02	Cikaba
ER0030	26-May-03	Unknown
ER0033	Uknown	Unknown

Sample locations are shown in Figure 4.2.

Table 4.2. Species-diagnostic nucleotide sites present in three nuclear genes examined in Eritrean elephants.

	BGN												*PHKA2*					*PLP*			
	N=	287	**304**	308	472	**485**	499	508	513	515	516	**570**	N=	**39**	71	**871**	872	N=	319	345	**361**
E. maximus	12	G	C	G	A	C	T	G	A	GD	D	T	12	C	G	C	G	12	G	C	G
L. cyclotis	116	GA	C	GC	AG	C	D	T	G	D	D	T	71	C	A	C	GA	118	GA	T	G
L. africana	806	G	**T**	G	A	**T**	DT	**G**	G	GD	DG	**C**	721	**T**	A	**T**	G	661	G	T	**A**
ER0001		G	**T**	G	A	**T**	T	**G**	G	G	G	**C**		--	--	--	--		--	--	--
ER0002		G	**T**	G	A	**T**	T	**G**	G	G	G	**C**		**T**	A	**T**	G		G	T	**A**
ER0003		G	**T**	G	A	**T**	T	**G**	G	G	G	**C**		**T**	A	**T**	G		G	T	**A**
ER0004		G	**T**	G	--	--	--	--	--	--	--	--		**T**	A	**T**	G		G	T	**A**
ER0007		G	**T**	G	A	**T**	T	**G**	G	G	G	**C**		**T**	A	**T**	G		G	T	**A**
ER0008		G	**T**	G	--	--	--	--	--	--	--	--		**T**	A	**T**	G		G	T	**A**
ER0022		G	**T**	G	A	**T**	T	**G**	G	G	G	**C**		**T**	A	**T**	G		G	T	**A**
ER0024		--	--	--	--	--	--	--	--	--	--	--		**T**	A	--	--		--	--	--
ER0025		--	--	--	--	--	--	--	--	--	--	--		**T**	A	--	--		--	--	--

Species diagnostic nucleotide sites in three genes that distinguish savanna (*L. africana*) from forest *(L. cyclotis*) elephants (Ishida et al. 2011a) are indicated in boldface; at variable sites both bases are shown. N is the number of chromosomes sequenced by previous trans-national studies for each of three species of elephant (Ishida et al. 2011a, Lei et al. 2009, Roca et al. 2001 and 2005); results are also shown for individual Eritrean elephants (code ER). Nucleotide site positions are from Roca et al. (2005). Dash indicates that sequence was not generated for a sample; D indicates a deletion.

Table 4.3. Fisher's exact tests comparing Eritrean sequences to those of forest or savanna elephants.

	BGN		*PHKA2*		*PLP*	
	L. cyclotis-typical sequences	*L. africana*-typical sequences	*L. cyclotis*-typical sequences	*L. africana*-typical sequences	*L. cyclotis*-typical sequences	*L. africana*-typical sequences
Eritrean elephants	0	7	0	8	0	6
Forest elephants	116	0	71	0	118	0
	P=	*0.0000*	*P=*	*0.0000*	*P=*	*0.0000*
Eritrean elephants	0	7	0	8	0	6
Savanna elephants	0	806	0	721	2	661
	P=	*1.0000*	*P=*	*1.0000*	*P=*	*0.9821*

Chromosome numbers are listed for each of three unlinked nuclear genes. Rows list the number of elephant chromosomes examined by the current study for elephant individuals from Eritrea vs. the number of chromosomes previously examined for either forest (*Loxodonta cyclotis*) or savanna (*L. africana*) elephants (Ishida *et al.* 2011a; Lei *et al.* 2009; Roca *et al.* 2005; Roca *et al.* 2001). Columns show the number of chromosomal sequences that matched those previously shown to be typical for *L. cyclotis* or *L. africana*. Sex of the Eritrean elephants was unknown; therefore, nuclear amplicons were conservatively estimated as representative of 1 rather than 2 X-chromosomes. For the *PLP* gene, *Loxodonta africana* includes 2 putative hybrid elephants from Cameroon (Roca *et al.* 2005).

Table 4.4. Calculations to examine the number of elephant population subdivisions.

Admixture - Correlated

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	3	-17251.6	0.458258	—	—	—
2	3	-14511.03333	0.763763	2740.566667	2546.9	3334.674868
3	3	-14317.36667	1.357694	193.666667	91.133333	67.123612
4	3	-14214.83333	1.154701	102.533333	43.5	37.672105
5	3	-14155.8	2.306513	59.033333	22.766667	9.870602
6	3	-14119.53333	2.83784	36.266667	35	12.333326
7	3	-14118.26667	6.017752	1.266667	19.4	3.223795
8	3	-14136.4	10.049378	-18.133333	8.766667	0.872359
9	3	-14145.76667	5.53022	-9.366667	0.966667	0.174797
10	3	-14154.16667	9.303942	-8.4	24.833333	2.66912
11	3	-14187.4	21.589118	-33.233333	39.3	1.820362
12	3	-14259.93333	16.519786	-72.533333	3.7	0.223974
13	3	-14328.76667	1.550269	-68.833333	5.533333	3.569274
14	3	-14403.13333	16.163642	-74.366667	7.733333	0.47844
15	3	-14469.76667	17.86962	-66.633333	8.4	0.470072
16	3	-14544.8	29.247051	-75.033333	21.666667	0.740815
17	3	-14641.5	57.161963	-96.7	10.533333	0.184272
18	3	-14727.66667	62.716532	-86.166667	137.433333	2.191341
19	3	-14951.26667	15.267067	-223.6	186.466667	12.213654
20	3	-14988.4	30.831153	-37.133333	67.866667	2.201237
21	3	-15093.4	119.902002	-105	95.8	0.798986
22	3	-15294.2	102.302346	-200.8	62.166667	0.607676
23	3	-15557.16667	71.396242	-262.966667	337.933333	4.733209
24	3	-15482.2	67.000224	74.966667	—	—

Table 4.4. Continued.

Admixture - Independent

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	3	-17252.4	0.1	—	—	—
2	3	-14598.26667	0.11547	2654.133333	2645.5	22910.70206
3	3	-14589.63333	1.361372	8.633333	83.033333	60.992397
4	3	-14664.03333	1.569501	-74.4	112.8	71.869977
5	3	-14625.63333	1.096966	38.4	84.733333	77.243389
6	3	-14671.96667	14.365352	-46.333333	4.766667	0.331817
7	3	-14713.53333	9.282421	-41.566667	11.633333	1.253265
8	3	-14743.46667	4.62313	-29.933333	2.1	0.454238
9	3	-14775.5	6.657327	-32.033333	5.933333	0.891249
10	3	-14801.6	7.808329	-26.1	11.533333	1.477055
11	3	-14839.23333	8.894005	-37.633333	22.466667	2.526046
12	3	-14854.4	22.440811	-15.166667	20.9	0.931339
13	3	-14890.46667	10.96464	-36.066667	0.5	0.045601
14	3	-14926.03333	2.173323	-35.566667	21.866667	10.061397
15	3	-14939.73333	6.439203	-13.7	24.366667	3.784112
16	3	-14977.8	10.278132	-38.066667	23.266667	2.263706
17	3	-14992.6	16.284041	-14.8	23.333333	1.432896
18	3	-15030.73333	3.286842	-38.133333	20.466667	6.226847
19	3	-15048.4	13.383946	-17.666667	19.5	1.456969
20	3	-15085.56667	3.350124	-37.166667	13	3.880453
21	3	-15109.73333	3.652853	-24.166667	5.466667	1.496547
22	3	-15128.43333	6.296295	-18.7	15.133333	2.40353
23	3	-15162.26667	17.339358	-33.833333	2.633333	0.15187
24	3	-15198.73333	9.469072	-36.466667	—	—

Table 4.4. Continued.

Noadmixture - Correlated

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	3	-17251.76667	0.152753	—	—	—
2	3	-14501.96667	0.305505	2749.8	2544.766667	8329.704197
3	3	-14296.93333	0.208167	205.033333	92.833333	445.956908
4	3	-14184.73333	0.51316	112.2	46	89.640633
5	3	-14118.53333	1.021437	66.2	20.3	19.873964
6	3	-14072.63333	0.416333	45.9	27.1	65.092095
7	3	-14053.83333	3.601851	18.8	14.266667	3.960926
8	3	-14049.3	4.340507	4.533333	0.933333	0.215029
9	3	-14045.7	3.315117	3.6	9.5	2.865661
10	3	-14051.6	3.835362	-5.9	22.833333	5.953371
11	3	-14080.33333	5.346338	-28.733333	26.466667	4.950429
12	3	-14135.53333	8.6031	-55.2	25	2.905929
13	3	-14215.73333	17.146525	-80.2	3.8	0.221619
14	3	-14299.73333	18.914104	-84	36.6	1.935064
15	3	-14420.33333	19.935981	-120.6	47.133333	2.364234
16	3	-14493.8	32.897264	-73.466667	72.766667	2.211937
17	3	-14640.03333	107.837115	-146.233333	65.866667	0.610798
18	3	-14720.4	23.84198	-80.366667	50.8	2.130696
19	3	-14851.56667	74.482638	-131.166667	102.833333	1.380635
20	3	-15085.56667	58.132636	-234	170.233333	2.928361
21	3	-15149.33333	114.366793	-63.766667	238.366667	2.08423
22	3	-15451.46667	161.061365	-302.133333	154.266667	0.957813
23	3	-15599.33333	154.876736	-147.866667	11.333333	0.073176
24	3	-15758.53333	294.328547	-159.2	—	—

The *ad hoc* method of Evanno et al. (2005) was used to examine the number of population subdivisions for elephants across Africa. The method was implemented in Structure Harvester (Earl & vonHoldt 2012). Calculations utilized the results from STRUCTURE software based on 1 million Markov chain Monte Carlo generations following a burn-in of 100,000 steps and 3 repetitions. Calculations for "Noadmixture - Independent" could not be performed since the standard deviation of the estimate of Ln Pr(Data) was less than 0.0000001. Highlighted rows indicate the most likely value of K.

LITERATURE CITED

Akaike H (1974) A new look at the statistical model identification. *Automatic Control, IEEE Transactions on* **19**, 716-723.

Anonymous (2013) Laelai-Gash sub-zone registering significant accomplishment regarding forestry and wildlife.

Archie EA, Hollister-Smith JA, Poole JH, *et al.* (2007) Behavioural inbreeding avoidance in wild African elephants. *Molecular Ecology* **16**, 4138-4148.

Archie EA, Maldonado JE, Hollister-Smith JA, *et al.* (2008) Fine-scale population genetic structure in a fission-fusion society. *Molecular Ecology* **17**, 2666-2679.

Archie EA, Moss CJ & Alberts SC (2003) Characterization of tetranucleotide microsatellite loci in the African savannah elephant (*Loxodonta africana africana*). *Molecular Ecology Notes* **3**, 244-246.

Arnason U, Adegoke JA, Gullberg A, *et al.* (2008) Mitogenomic relationships of placental mammals and molecular estimates of their divergences. *Gene* **421**, 37-51.

Arnason U, Gullberg A & Janke A (1999) The mitochondrial DNA molecule of the aardvark, *Orycteropus afer*, and the position of the Tubulidentata in the eutherian tree. *Proceedings of the Royal Society of London B Biological Sciences* **266**, 339-345.

Baker CS (2013) Journal of Heredity Adopts Joint Data Archiving Policy. *Journal of Heredity* **104**, 1-1.

Balloux F & Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11**, 155-165.

Bandelt HJ, Forster P & Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**, 37-48.

Barnes I, Shapiro B, Lister A, *et al.* (2007) Genetic structure and extinction of the woolly mammoth, *Mammuthus primigenius*. *Current Biology* **17**, 1072-1075.

Barnes RFW (1999) Is there a future for elephants in West Africa? *Mammal Review* **29**, 175-199.

Barnes RFW, Craig GC, Dublin HT, *et al.* (1999) African elephant database 1998. *Occasional Papers of the IUCN Species Survival Commission* **22**, i-vi, 1-249.

Barriel V, Thuet E & Tassy P (1999) Molecular phylogeny of Elephantidae. Extreme divergence of the extant forest African elephant. *Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie* **322**, 447-454.

Belkhir K, Borsa P, Chikhi L, Raufaste N & Bonhomme F (2004) Genetix 4.05. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier.

Blanc JJ, Barnes RFW, Craig GC, *et al.* (2007) African elephant status report 2007. An update from the African elephant database. *Occasional Papers of the IUCN Species Survival Commission* **33**, i-vi, 1-275.

Bonnefille R (2010) Cenozoic vegetation, climate changes and hominid evolution in tropical Africa. *Global and Planetary Change* **72**, 390-411.

Boutin-Ganache I, Raposo M, Raymond M & Deschepper CF (2001) M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. *Biotechniques* **31**, 24-26, 28.

Brandt AL, Hagos Y, Yacob Y, *et al.* (2014) The elephants of Gash-Barka, Eritrea: Nuclear and mitochondrial genetic patterns. *Journal of Heredity*, In Press.

Brandt AL, Ishida Y, Georgiadis NJ & Roca AL (2012) Forest elephant mitochondrial genomes reveal that elephantid diversification in Africa tracked climate transitions. *Molecular Ecology* **21**, 1175-1189.

Capelli C, MacPhee RD, Roca AL, *et al.* (2006) A nuclear DNA phylogeny of the woolly mammoth (*Mammuthus primigenius*). *Molecular Phylogenetics and Evolution* **40**, 620-627.

Cerling TE, Ehleringer JR & Harris JM (1998) Carbon dioxide starvation, the development of C-4 ecosystems, and mammalian evolution. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **353**, 159-170.

Cerling TE, Harris JM, MacFadden BJ, *et al.* (1997) Global vegetation change through the Miocene/Pliocene boundary. *Nature* **389**, 153-158.

Comstock KE, Georgiadis N, Pecon-Slattery J, *et al.* (2002) Patterns of molecular genetic variation among African elephant populations. *Molecular Ecology* **11**, 2489-2498.

Cornuet JM & Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001-2014.

Crispo E & Hendry AP (2005) Does time since colonization influence isolation by distance? A meta-analysis. *Conservation Genetics* **6**, 665-682.

Currat M, Ruedi M, Petit RJ & Excoffier L (2008) The hidden side of invasions: massive introgression by local genes. *Evolution* **62**, 1908-1920.

Debruyne R (2005) A case study of apparent conflict between molecular phylogenies: the interrelationships of African elephants. *Cladistics* **21**, 31-50.

Debruyne R, Barriel V & Tassy P (2003a) Mitochondrial *cytochrome b* of the Lyakhov mammoth (Proboscidea, Mammalia): new data and phylogenetic analyses of Elephantidae. *Molecular Phylogenetics and Evolution* **26**, 421-434.

Debruyne R, Chu G, King CE, *et al.* (2008) Out of America: ancient DNA evidence for a new world origin of late quaternary woolly mammoths. *Current Biology* **18**, 1320-1326.

Debruyne R, Van Holt A, Barriel V & Tassy P (2003b) Status of the so-called African pygmy elephant (*Loxodonta pumilio* [NOACK 1906]): phylogeny of *cytochrome b* and mitochondrial control region sequences. *Comptes Rendus Biologies* **326**, 687-697.

DeMenocal PB (2004) African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth and Planetary Science Letters* **220**, 3-24.

Deraniyagala PEP (1955) *Some extinct elephants, their relatives and the two living species* Government Press (National Museum of Ceylon), Colombo.

Dirienzo A, Peterson AC, Garza JC, *et al.* (1994) Mutational processes of simple-sequence repeat loci in human-populations. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 3166-3170.

Douglas-Hamilton I (1987) African elephants: population trends and their causes. *Oryx* **21**, 11-24.

Drummond A, Ashton B, Buxton S, *et al.* (2010) Geneious v5.1, Available from http://www.geneious.com.

Drummond AJ, Ho SYW, Phillips MJ & Rambaut A (2006) Relaxed Phylogenetics and Dating with Confidence. *PLOS Biology* **4**, e88.

Drummond AJ & Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**, 214.

Earl D & vonHoldt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**, 359-361.

Eggert LS, Rasner CA & Woodruff DS (2002) The evolution and phylogeography of the African elephant inferred from mitochondrial DNA sequence and nuclear microsatellite markers. *Proceedings of the Royal Society of London B Biological Sciences* **269**, 1993-2006.

Enk J, Devault A, Debruyne R, *et al.* (2011) Complete Columbian mammoth mitogenome suggests interbreeding with woolly mammoths. *Genome Biology* **12**, R51.

Evanno G, Regnaut S & Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.

Excoffier L & Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564-567.

Fernando P, Pfrender ME, Encalada SE & Lande R (2000) Mitochondrial DNA variation, phylogeography and population structure of the Asian elephant. *Heredity* **84**, 362-372.

Fernando P, Vidya TN, Payne J, *et al.* (2003) DNA analysis indicates that Asian elephants are native to Borneo and are therefore a high priority for conservation. *PLOS Biology* **1**, E6.

Fernando PJ, Vidya TNC & Melnick DJ (2001) Isolation and characterization of tri- and tetranucleotide microsatellite loci in the Asian elephant, *Elephas maximus*. *Molecular Ecology Notes* **1**, 232-233.

Fleischer RC, Perry EA, Muralidharan K, Stevens EE & Wemmer CM (2001) Phylogeography of the Asian elephant *(Elephas maximus)* based on mitochondrial DNA. *Evolution* **55**, 1882-1892.

Folinsbee KE & Brooks DR (2007) Miocene hominoid biogeography: pulses of dispersal and differentiation. *Journal of Biogeography* **34**, 383-397.

Forest Fl (2009) Calibrating the Tree of Life: fossils, molecules and evolutionary timescales. *Annals of Botany* **104**, 789-794.

Georgiadis N, Bischof L, Templeton A, *et al.* (1994) Structure and history of African elephant populations: I. Eastern and southern Africa. *Journal of Heredity* **85**, 100-104.

Gibbard PL & Head MJ (2010) The newly-ratified definition of the Quaternary System/Period and redefinition of the Pleistocene Series/Epoch, and comparison of proposals advanced prior to formal ratification. *Episodes* **33**, 152-158.

Gilbert MT, Drautz DI, Lesk AM, *et al.* (2008) Intraspecific phylogenetic analysis of Siberian woolly mammoths using complete mitochondrial genomes. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 8327-8332.

Gilbert MT, Tomsho LP, Rendulic S, *et al.* (2007) Whole-genome shotgun sequencing of mitochondria from ancient hair shafts. *Science* **317**, 1927-1930.

Good JM, Hird S, Reid N, *et al.* (2008) Ancient hybridization and mitochondrial capture between two species of chipmunks. *Molecular Ecology* **17**, 1313-1327.

Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.

Gowers W (1948) African Elephants and Ancient Authors. *African Affairs London* **188**, pp. 173-180.

Griffin DL (2002) Aridity and humidity: two aspects of the late Miocene climate of North Africa and the Mediterranean. *Palaeogeography Palaeoclimatology Palaeoecology* **182**, 65-91.

Groves CP & Grubb P (2000) Do *Loxodonta cyclotis* and *L. africana* interbreed? *Elephant* **2**, 4-7.

Grubb P, Groves CP, Dudley JP & Shoshani J (2000) Living African elephants belong to two species: *Loxodonta africana* (Blumenbach, 1797) and *Loxodonta cyclotis* (Matschie, 1900). *Elephant* **2**, 1-4.

Hagos Y, Yacob Y, Ghebrehiwet M & Shoshani J (2003) The elephants (*Loxodonta africana*) of Gash-Barka, Eritrea: Part. 1. historical perspective and related findings. *Pachyderm* **34**, 13-23.

Hailer F, Kutschera VE, Hallstroem BM, *et al.* (2012) Nuclear Genomic Sequences Reveal that Polar Bears Are an Old and Distinct Bear Lineage. *Science* **336**, 344-347.

Hanke M & Wink M (1994) Direct DNA sequencing of PCR-amplified vector inserts following enzymatic degradation of primer and dNTPs. *Biotechniques* **17**, 858-860.

Hartl DL & Clark AG (1997) *Principles of population genetics, Third edition*.

Hauf J, Waddell PJ, Chalwatzis N, Joger U & Zimmermann FK (1999) The complete mitochondrial genome sequence of the African elephant (*Loxodonta africana*),

phylogenetic relationships of Proboscidea to other mammals, and D-loop heteroplasmy. *Zoology-Analysis of Complex Systems* **102**, 184-195.

Hawks J, Hunley K, Lee SH & Wolpoff M (2000) Population bottlenecks and Pleistocene human evolution. *Molecular Biology and Evolution* **17**, 2-22.

Haynes G (1991) *Mammoths, mastodonts, and elephants : biology, behavior, and the fossil record* Cambridge University Press, Cambridge New York.

Hedrick PW (2007) Sex: differences in mutation, recombination, selection, gene flow, and genetic drift. *Evolution* **61**, 2750-2771.

Hoelzer GA (1997) Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees revisited. *Evolution* **51**, 622-626.

Hollister-Smith JA, Poole JH, Archie EA, *et al.* (2007) Age, musth and paternity success in wild male African elephants, *Loxodonta africana*. *Animal Behaviour* **74**, 287-296.

Hubisz MJ, Falush D, Stephens M & Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* **9**, 1322-1332.

Ingman M, Kaessmann H, Paabo S & Gyllensten U (2000) Mitochondrial genome variation and the origin of modern humans. *Nature* **408**, 708-713.

Ishida Y, Demeke Y, de Groot PJV, *et al.* (2011a) Distinguishing forest and savanna African elephants using short nuclear DNA sequences. *Journal of Heredity* **102**, 610-616.

Ishida Y, Georgiadis NJ, Hondo T & Roca AL (2013) Triangulating the provenance of African elephants using mitochondrial DNA. *Evolutionary Applications* **6**, 253-265.

Ishida Y, Oleksyk TK, Georgiadis NJ, *et al.* (2011b) Reconciling apparent conflicts between mitochondrial and nuclear phylogenies in African elephants. *Plos One* **6**, 16.

Jensen JL, Bohonak AJ & Kelley ST (2005) Isolation by distance, web service. *Bmc Genetics* **6**.

Johnson MB, Clifford SL, Goossens B, *et al.* (2007) Complex phylogeographic history of central African forest elephants and its implications for taxonomy. *BMC Evol Biol* **7**, 244.

Karesh W, Smith F & Frazier-Taylor H (1989) A remote method for obtaining skin biopsy samples. *Conservation Biology* **1**, 261-262.

Kingdon J (1979) *East African mammals: an atlas of evolution in Africa. Volume III Part B (Large mammals)* Academic Press, London.

Krause J, Dear PH, Pollack JL, *et al.* (2006) Multiplex amplification of the mammoth mitochondrial genome and the evolution of Elephantidae. *Nature* **439**, 724-727.

Largen MJ & Yalden DW (1987) The decline of elephant and black rhinoceros in Ethiopia. *Oryx* **21**, 103-106.

Larkin MA, Blackshields G, Brown NP, *et al.* (2007) Clustal W and clustal X version 2.0. *Bioinformatics* **23**, 2947-2948.

Leakey MG & Harris JM (2003) *Lothagam : the dawn of humanity in eastern Africa* Columbia University Press, New York.

Lei R, Brenneman RA & Louis EE (2008) Genetic diversity in the North American captive African elephant collection. *Journal of Zoology* **275**, 252-267.

Lei R, Brenneman RA, Schmitt DL & Louis EE, Jr. (2009) Detection of cytonuclear genomic dissociation in the North American captive African elephant collection. *Journal of Heredity* **100**, 675-680.

Lei R, Brenneman RA, Schmitt DL & Louis EE, Jr. (2012) Genetic diversity in North American captive Asian elephants. *Journal of Zoology* **286**, 38-47.

Leuenberger H (1955) *Aethiopien, Kaiserreich seit Salomon*, Stauffacher-Verlag, Zurich.

Levene H (1949) On a matching problem arising in genetics. *Annals of Mathematical Statistics* **20**, 91-94.

Lindsey PA, Alexander R, Mills MGL, Romanach S & Woodroffe R (2007) Wildlife viewing preferences of visitors to protected areas in South Africa: implications for the role of ecotourism in conservation. *Journal of Ecotourism* **6**, 19-33.

Lisiecki LE & Raymo ME (2005) A Pliocene-Pleistocene stack of 57 globally distributed benthic delta O-18 records. *Paleoceanography* **20**, 17.

Litoroh MW (1997) Aerial census of the Gash-Setit elephant population of Eritrea and Ethiopia. *Pachyderm* **23**, 12-18.

Lowe AJ, Harris D, Dormontt E & Dawson IK (2010) Testing putative African tropical forest refugia using chloroplast and nuclear DNA phylogeography. *Tropical Plant Biol* **3**, 50-58.

Maglio VJ (1973) Origin and evolution of the Elephantidae. *Trans. Am. Phil. Soc. Philad., New Series* **63**, 1-149.

Maikaew U, Wajjwalku W, Siripholwat V, *et al.* (2007) The study of nucleotides arrangement of mitochondrial DNA in Thai elephants. *KKU Vet. J.* **17**, 11-21.

Maley J (2001) The impact of arid phases on the African rain forest through geological history. In: *African Rain Forest Ecology and Conservation: An Interdisciplinary Perspective* (eds. Weber W, White LJT, Vedder A, Naughton-Treves L), pp. xii, 588 p. Yale University Press, New Haven [Conn.].

Mayr E (1942) *Systematics and the origin of species* (Columbia Univ. Press).

Mayr E (1963) *Animal species and evolution* Harvard University Press, Cambridge, MA.

Mayr E & Ohara RJ (1986) The Biogeographic Evidence Supporting the Pleistocene Forest Refuge Hypothesis. *Evolution* **40**, 55-67.

McKenzie M, Chiotis M, Pinkert CA & Trounce IA (2003) Functional respiratory chain analyses in murid xenomitochondrial cybrids expose coevolutionary constraints of cytochrome b and nuclear subunits of complex III. *Molecular Biology and Evolution* **20**, 1117-1124.

McWilliam H, Valentin F, Goujon M, *et al.* (2009) Web services at the European Bioinformatics Institute-2009. *Nucleic Acids Research* **37**, W6-W10.

Meier R & Wheeler Q (2000) *Species concepts and phylogenetic theory: A debate* Columbia University Press, New York.

Menotti-Raymond MA, David VA, Wachter LL, Butler JM & O'Brien SJ (2005) An STR forensic typing system for genetic individualization of domestic cat (*Felis catus*) samples. *J Forensic Sci* **50**, 1061-1070.

Miller W, Drautz DI, Ratan A, *et al.* (2008) Sequencing the nuclear genome of the extinct woolly mammoth. *Nature* **456**, 387-390.

Morin PA, Manaster C, Mesnick SL & Holland R (2009) Normalization and binning of historical and multi-source microsatellite data: overcoming the problems of allele size shift with allelogram. *Molecular Ecology Resources* **9**, 1451-1455.

Murata Y, Yonezawa T, Kihara I, *et al.* (2009) Chronology of the extant African elephant species and case study of the species identification of the small African elephant with the molecular phylogenetic method. *Gene* **441**, 176-186.

Murphy WJ & O'Brien SJ (2007) Designing and optimizing comparative anchor primers for comparative gene mapping and phylogenetic inference. *Nature Protocols* **2**, 3022-3030.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583-590.

Nyakaana S & Arctander P (1999) Population genetic structure of the African elephant in Uganda based on variation at mitochondrial and nuclear loci: evidence for male-biased gene flow. *Molecular Ecology* **8**, 1105-1115.

Nyakaana S, Arctander P & Siegismund HR (2002) Population structure of the African savannah elephant inferred from mitochondrial control region sequences and nuclear microsatellite loci. *Heredity* **89**, 90-98.

O'Connell-Rodwell CE, Rodwell T, Rice M & Hart LA (2000) Living with the modern conservation paradigm: can agricultural communities co-exist with elephants? A five-year case study in East Caprivi, Namibia. *Biological Conservation* **93**, 381-391.

Ohta T (1973) Slightly deleterious mutant substitutions in evolution. *Nature* **246**, 96-98.

Okello JB, Masembe C, Rasmussen HB, *et al.* (2008) Population Genetic Structure of Savannah Elephants in Kenya: Conservation and Management Implications. *Journal of Heredity*.

Otero O, Lecuyer C, Fourel F, *et al.* (2011) Freshwater fish delta O-18 indicates a Messinian change of the precipitation regime in Central Africa. *Geology* **39**, 435-438.

Peakall R & Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**, 2537-2539.

Petit RJ & Excoffier L (2009) Gene flow and species delimitation. *Trends Ecol Evol* **24**, 386-393.

Polybius (1923) *The Histories: Volume III: Books 5-8* Harvard University Press.

Poole JH (1989a) Mate guarding, reproductive success and female choice in African elephants. *Animal Behaviour* **37**, 842-849.

Poole JH (1989b) Announcing intent: the aggressive state of musth in African elephants. *Animal Behaviour* **37**, 140-152.

Poole JH, Lee PC, Nijraini N & Moss CJ (2011) *Longevity, competition, and musth: a long-term perspective on male reproductive strategies* The University of Chicago Press.

Posada D & Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817-818.

Pritchard JK, Stephens M & Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Rasmussen DT & Gutierrez M (2009) A mammalian fauna from the Late Oligocene of northwestern Kenya. *Palaeontographica Abteilung a-Palaozoologie-Stratigraphie* **288**, 1-52.

Rasmussen HB, Okello JBA, Wittemyer G, *et al.* (2008) Age- and tactic-related paternity success in male African elephants. *Behavioral Ecology* **19**, 9-15.

Raymond M & Rousset F (1995) Genepop (Version-1.2) - Population-genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.

Reed DH & Frankham R (2003) Correlation between fitness and genetic diversity. *Conservation Biology* **17**, 230-237.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution* **43**, 223-225.

Roca AL (2008) The mastodon mitochondrial genome: a mammoth accomplishment. *Trends in Genetics* **24**, 49-52.

Roca AL, Geogiadis N & O'Brien SJ (2002) African elephant genetics: request for samples. *Pachyderm* **33**, 93-95.

Roca AL, Georgiadis N & O'Brien SJ (2005) Cytonuclear genomic dissociation in African elephant species. *Nature Genetics* **37**, 96-100.

Roca AL, Georgiadis N & O'Brien SJ (2007) Cyto-nuclear genomic dissociation and the African elephant species question. *Quaternary International* **169-170**, 4-16.

Roca AL, Georgiadis N, Pecon-Slattery J & O'Brien SJ (2001) Genetic evidence for two species of elephant in Africa. *Science* **293**, 1473-1477.

Roca AL & O'Brien SJ (2005) Genomic inferences from Afrotheria and the evolution of elephants. *Current Opinion in Genetics & Development* **15**, 652-659.

Rogaev EI, Moliaka YK, Malyarchuk BA, *et al.* (2006) Complete mitochondrial genome and phylogeny of Pleistocene mammoth *Mammuthus primigenius*. *PLOS Biology* **4**, e73.

Rohland N, Malaspinas AS, Pollack JL, *et al.* (2007) Proboscidean mitogenomics: chronology and mode of elephant evolution using mastodon as outgroup. *PLOS Biology* **5**, e207.

Rohland N, Reich D, Mallick S, *et al.* (2010) Genomic DNA sequences from mastodon and woolly mammoth reveal deep speciation of forest and savanna elephants. *PLOS Biology* **8**, e1000564.

Roth HH & Douglas-Hamilton I (1991) Distribution and status of elephants in west Africa. *Mammalia* **55**, 489-527.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**, 1219-1228.

Rousset F (2008) GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.

Salzmann U & Hoelzmann P (2005) The Dahomey Gap: an abrupt climatically induced rain forest fragmentation in West Africa during the late Holocene. *Holocene* **15**, 190-199.

Sambrook J, Fritsch EF & Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sanders WJ, Gheerbrant E, Harris JM, Saegusa H & Delmer C (2010) Proboscidea. In: *Cenozoic Mammals of Africa* (eds. Werdelin L, Sanders WJ). University of California Press, Berkeley.

Sepulchre P, Ramstein G, Fluteau F, *et al.* (2006) Tectonic uplift and Eastern Africa aridification. *Science* **313**, 1419-1423.

Shoshani J, Hagos Y & Yacob Y (2000) Observations on elephant habitat and conservation of elephants in Eritrea. *Elephant* **2**, 14-19.

Shoshani J, Hagos Y, Yacob Y, Ghebrehiwet M & Kebrom E (2004) Elephants (*Loxodonta africana*) of Zoba Gash Barka, Eritrea: Part 2. Numbers and distribution, ecology and behaviour and fauna and flora in their ecosystem. *Pachyderm* **36**, 52-68.

Shoshani J & Tassy P (1996) *The Proboscidea: Evolution and Palaeoecology of Elephants and their Relatives* Oxford University Press, Oxford ; New York.

Slatkin M (1993) Isolation by distance in equilibrium and nonequilibrium populations. *Evolution* **47**, 264-279.

Sukumar R (1991) The management of large mammals in relation to male strategies and conflict with people. *Biological Conservation* **55**, 93-102.

Swofford DL (2002) *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10.* Sinauer, Sunderland, MA.

Thouless CR (1995) Long distance movements of elephants in northern Kenya. *African Journal of Ecology* **33**, 321-334.

Vidya TN, Sukumar R & Melnick DJ (2009) Range-wide mtDNA phylogeography yields insights into the origins of Asian elephants. *Proceedings of the Roual Society B-Biological Sciences* **276**, 893-902.

Vignaud P, Duringer P, Mackaye HT, *et al.* (2002) Geology and palaeontology of the Upper Miocene Toros-Menalla hominid locality, Chad. *Nature* **418**, 152-155.

Wasser SK, Clark B & Laurie C (2009) The ivory trail. *Sci Am* **301**, 68-74, 76.

Weckerly FW (1998) Sexual-size dimorphism: Influence of mass and mating systems in the most dimorphic mammals. *Journal of Mammalogy* **79**, 33-52.

Weldeyohannes M & Siratu TG (2010) *The State of Eritrea. Ministry of Land, Water and Environment. Department of Environment. The 4th National Report to the Convention on Biological Diversity* Asmara, Eritrea.

Werdelin L & Sanders WJ (2010) *Cenozoic Mammals of Africa* University of California Press, Berkeley.

White F (1983) *The Vegetation of Africa* UNESCO, Paris.

Wilson GA & Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**, 1177-1191.

Wittemyer G, Daballen D & Douglas-Hamilton I (2013) Comparative Demography of an At-Risk African Elephant Population. *Plos One* **8**.

Wittemyer G, Northrup JM, Blanc J, *et al.* (2014) Illegal killing for ivory drives global decline in African elephants. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 13117-13121.

Woolley L-A, Mackey RL, Page BR & Slotow R (2008) Modelling the effect of age-specific mortality on elephant *Loxodonta africana* populations: can natural mortality provide regulation? *Oryx* **42**, 49-57.

Wright S (1938) Size of population and breeding structure in relation to evolution. *American Association for the Advancement of Science. Science* **87**, 430-431.

Wright S (1943) Isolation by distance. *Genetics* **28**, 114-138.

Wright S (1978) *Evolution and the genetics of populations, variability within and among natural populations*.

Yacob Y, Shoshani J, Hagos Y & Kebrom E (2004) The elephants (*Loxodonta africana*) of Zoba Gash Barka, Eritrea: part 3. Ecological and other data from tusks, teeth and carcasses. *Pachyderm* **37**, 45-58.

Yalden DW, Largen MJ & Kock D (1986) Catalog of the mammals of Ethiopia .6. Perissodactyla, Proboscidea, Hyracoidea, Lagomorpha, Tubulidentata, Sirenia and Cetacea. *Monitore Zoologico Italiano-Italian Journal of Zoology* **21**, 31-103.

Yeh FC & Boyle TJB (1997) Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian Journal of Botany* **129**, 157-157.

APPENDIX A: AFRICAN ELEPHANT GENOTYPE DATA

ID	Pop	LAF13	LAF11	EMX4	LaT05	LaT06	LAF10	LAF12	EMX5	EMX3	LAF29	LAF37
SL0001	1	000000	247247	369369	000000	379395	162166	000000	000000	000000	000000	215219
LO3501	2	294298	247247	369369	357357	367383	166166	210214	269269	264264	207207	217223
LO3502	2	310318	247247	369369	385385	367399	162162	214214	265269	261267	205211	219229
LO3503	2	294313	247247	369369	477498	351375	166166	214214	265269	264264	209209	215215
LO3505	2	294310	247247	369369	389389	327387	166166	206214	265269	264264	203213	215223
LO3507	2	306313	247247	369369	505509	379395	162162	210214	265269	264264	213213	215225
LO3508	2	294310	247247	369369	357357	359399	162166	206214	269269	264264	211211	215217
LO3509	2	302310	247247	369369	498521	367379	166166	194218	269269	264267	209219	215227
LO3510	2	306310	247247	353369	477480	323359	162162	206238	269269	261264	207209	215217
LO3511	2	298306	247247	369373	389393	327375	162162	210214	269269	264267	209209	217225
LO3512	2	294302	247247	369369	289293	347407	162166	206214	265269	264267	209215	217217
LO3513	2	313326	247247	369369	389389	403407	162166	210214	269269	264267	209213	215217
LO3514	2	298318	247247	369369	364385	367375	162162	214230	265269	261264	205207	215229
LO3515	2	298318	247247	369369	360360	395399	162162	214214	265265	264267	209213	207223
LO3516	2	298302	247247	369369	488498	395399	162162	190190	265265	261267	209213	207217
LO3517	2	298330	247247	369369	357357	367379	166166	214214	269269	264264	211211	217217
LO3522	2	306310	247247	369373	000000	323395	162162	206210	265265	264267	209209	215223
OD0001	3	298310	247247	369369	477505	367399	162162	206214	269269	264264	215217	215223
DS1501	4	282294	247247	369369	281289	351391	162162	210214	265265	261264	209215	223225
DS1502	4	298313	247247	369369	277285	367387	166166	210226	269269	264264	205211	215223
DS1503	4	302310	247247	369369	285285	327407	162162	214222	269269	264267	209211	215223
DS1504	4	310318	247247	369373	441445	379407	162166	198210	265265	264264	203217	215223
DS1505	4	226322	247247	369369	277293	351403	162166	202210	269269	264264	209211	217223
DS1506	4	318322	247247	369369	465473	399403	162162	000000	269269	264264	211211	215215
DS1507	4	306310	247247	353369	360360	363379	162162	226238	269269	264264	203209	215217
DS1508	4	000000	247247	369369	493498	399399	162162	226226	269269	264267	203211	207217
DS1509	4	000000	243247	369373	374374	343379	162162	202202	269269	264267	205211	213217
DS1510	4	318322	247247	369369	364364	391399	162162	218218	269269	264267	209209	223223
DS1511	4	302318	247247	369369	437437	391399	166166	206218	269269	264264	209213	215215
DS1512	4	298310	243243	369369	379397	395399	162166	214214	265269	264264	209209	215221
DS1513	4	226310	247247	369373	469484	379411	162166	198214	265269	264306	209209	215225
DS1514	4	302318	247247	361369	305305	363379	162166	214214	265269	264267	203207	223223
DS1515	4	313318	247247	369369	437441	315347	162162	000000	269269	264264	205211	213223
DS1516	4	000000	247247	361361	360360	371371	162166	202206	269269	264264	203211	221223
DS1518	4	302318	247247	369369	364364	315391	162162	218218	269269	264264	209211	213223
DS1519	4	226310	247247	361369	285289	383383	166166	202214	265269	264267	211223	215215
DS1520	4	306318	247247	369373	364364	315379	162166	210210	269269	264264	203209	223223
DS1521	4	302302	247247	369373	345345	327375	162162	206218	265269	264267	203203	217217
DS1522	4	306313	243247	353369	360493	387387	162162	202210	269269	000000	000000	217225
DS1523	4	302313	247247	369377	281289	379387	162162	198198	269269	267267	205221	215225
DS1524	4	306310	247247	369369	498498	387391	162162	202206	265269	264267	211221	219225
DS1525	4	310310	243247	369373	364368	379407	162162	206206	265269	264264	209209	213221
DS1526	4	226306	247247	369369	393393	403419	162162	230230	269269	264264	215215	223223
DS1527	4	302310	247247	369369	480484	323407	162162	214222	269269	264267	211211	215223
DS1528	4	302310	243247	361369	469484	387399	162162	210218	269269	264264	215215	215223
DS1530	4	302313	247251	369369	371371	375379	162166	202210	269269	261267	209209	213215

ID	Pop	LAF13	LAF11	EMX4	LaT05	LaT06	LAF10	LAF12	EMX5	EMX3	LAF29	LAF37
DS1531	4	306306	247247	369369	477477	351391	166166	186210	265269	000000	209209	223223
DS1532	4	310310	243247	369369	353353	327363	162162	206210	265269	264264	209209	213223
DS1533	4	310313	247247	369369	277293	379403	162162	210226	269269	267267	211211	215217
DS1534	4	302322	247247	369369	349349	371403	166166	206214	269269	264264	215215	215215
DS1535	4	294318	247247	369369	477477	383399	166166	210214	269269	264264	203213	217217
DS1536	4	286330	243247	369369	433469	327387	162162	210226	265269	264267	203209	217221
DS1537	4	310313	247247	369373	477480	375379	162162	194218	269269	264264	207223	215217
DS1539	4	310318	247247	369369	357357	367387	162162	210226	269269	264264	203211	221221
DS1540	4	298322	247247	369369	389389	367411	162166	214218	269269	264264	209215	215217
DS1541	4	294313	247247	369369	449461	323355	162162	194210	265265	264264	209213	215217
DS1542	4	226318	247247	369369	364364	387391	162166	206226	265269	264264	203217	215223
DS1543	4	318318	247247	369369	364389	375407	166166	214226	269269	261264	203221	211223
DS1544	4	310326	247247	369369	360397	351399	162166	214218	269269	264306	197197	213223
DS1545	4	226313	247247	369369	465480	387395	162162	000000	269269	267267	209209	213223
DS1548	4	306330	243247	369373	493493	383411	162162	202202	269269	264306	203211	187217
DS1549	4	294294	247251	369369	484493	375379	158166	206226	265269	264267	203203	223237
DS1551	4	298306	247247	369369	473498	387395	162162	206206	265269	264264	211211	215215
DS1552	4	000000	247247	369373	357357	379387	166166	000000	000000	000000	209209	213215
DS1553	4	282306	247247	369369	465480	323379	162162	222222	265269	264267	203215	217225
DS1554	4	302306	247247	369369	349349	371403	162162	198206	269269	264264	211211	185215
DS1555	4	302318	247247	369369	488498	387419	162162	202210	265269	264264	203211	217221
DS1556	4	282318	243247	369369	360364	375379	162162	000000	269269	264264	209223	215223
DS1557	4	306310	247251	361369	488493	375407	162166	206210	265265	264267	209209	215223
DS1559	4	286302	247247	369377	493493	371379	162162	194202	265265	264306	209209	215217
DS1560	4	306313	247247	369377	368368	379395	166166	202210	265265	264264	203209	215217
NC0008	5	296298	247247	369369	461479	363383	162162	000000	000000	000000	209209	213213
NC0009	5	238302	247247	369369	349349	367375	162166	000000	000000	000000	213217	215223
NC0010	5	238302	247247	369369	345349	367375	162166	000000	000000	000000	213217	215223
NC0014	5	298310	247247	361369	445445	000000	162166	000000	000000	000000	205211	215217
GR0007	6	302310	247247	369377	313313	347387	000000	178206	265269	267267	209213	213217
GR0013	6	238306	247247	369377	313313	387407	162166	202210	265269	264267	205211	215217
GR0015	6	302306	247247	369369	480480	375375	162162	000000	265269	267267	215215	219219
GR0016	6	294302	247247	369369	493505	375399	162162	194210	265269	264264	215217	215215
GR0018	6	238282	247247	369377	277277	327411	162166	202202	265269	264264	213213	219219
GR0020	6	238298	247247	377377	405405	375391	166166	000000	265269	267267	211223	217217
GR0021	6	294306	247247	377377	401401	299379	000000	210214	265269	264267	209217	205223
GR0022	6	298313	247247	353369	488505	363367	162166	210218	269269	264264	213213	213213
GR0023	6	306313	247251	369369	368379	304383	162162	000000	265269	267267	211217	211215
GR0025	6	294318	247247	361369	357371	371395	162166	000000	269269	264264	215215	213221
GR0026	6	306326	247247	369369	445493	363379	162162	206206	265269	264267	209217	213213
GR0035	6	302330	247247	361377	313313	383387	162162	190206	265269	264267	209213	213215
GR0036	6	282310	247247	369377	371371	367371	162166	206206	269269	264267	203209	213215
GR0037	6	286318	247251	369377	277277	379399	162162	182206	269269	267267	207211	213215
GR0038	6	290302	247251	361369	309309	359387	162166	190206	265269	264267	211211	213215
GR0039	6	302318	247247	361377	353353	375403	162162	198214	269269	000000	211211	213217
GR0040	6	226310	247247	361369	493493	327355	162162	202206	265269	000000	205215	213219
GR0041	6	294310	247247	369373	453457	367399	166166	182202	269269	264264	213215	213225
GR0042	6	302302	247247	377377	457461	331399	162166	000000	265269	264267	215217	213217
BE4035	7	226238	251251	377377	281289	371383	166166	162162	269269	267267	209209	185217

ID	Pop	LAF13	LAF11	EMX4	LaT05	LaT06	LAF10	LAF12	EMX5	EMX3	LAF29	LAF37
BE4036	7	226226	251251	369377	301421	304375	166166	154182	269269	267267	211215	207217
BE4037	7	226238	251251	369377	289297	304304	158166	174182	269269	267267	217219	205205
BE4053	7	226238	251251	000000	277281	308383	158158	178190	269269	267267	217217	207217
BE4055	7	226238	251251	369377	281301	299383	158166	174182	269269	267267	217217	211217
BE4056	7	238238	251251	369377	293317	299387	158158	174178	269269	267267	217217	217217
BE4057	7	226238	251251	377377	289301	299367	158158	174194	269269	267267	217217	205211
BE4058	7	226238	251251	377377	297309	299375	158166	178186	269269	267267	217217	211217
BE4059	7	226294	251251	369377	289309	387391	166166	166186	269269	267267	215219	205217
WA4003	8	238238	251251	377377	285321	304383	158166	178194	269269	267267	217217	211219
WA4004	8	226226	251251	369377	305309	391395	158166	186186	269269	267267	211217	205211
WA4006	8	226238	251251	377377	285317	304387	158166	182190	269269	267267	217217	211217
WA4008	8	226238	251251	369377	277305	351391	158166	182182	269269	267267	211217	205219
WA4009	8	226238	251251	369369	273289	304304	158158	182190	269269	267267	217217	205217
WA4012	8	238322	251251	369369	285309	304379	158166	178194	269269	267267	217217	211217
WA4014	8	226238	251251	369377	285305	304308	166166	178182	269269	267267	213217	211217
WA4015	8	238238	251251	377377	297309	304383	158158	178186	269269	267267	217217	205205
WA4017	8	226238	251251	369377	281309	304383	166166	174186	269269	267267	217217	211217
WA4018	8	226238	247251	369377	285305	403403	166166	178178	269269	267267	217217	211219
WA4020	8	226238	251251	373377	277293	304308	166166	178186	269269	267267	217217	205211
WA4021	8	226238	251251	369377	281297	375379	158158	182186	269269	267267	217219	205205
WA4022	8	238318	251251	377377	269309	299304	158158	174178	269269	267267	217217	211217
WA4023	8	226238	251251	377377	289309	308375	158158	166186	269269	267267	211217	205205
WA4024	8	226238	251251	369377	281285	304383	158158	186194	269269	267267	217217	205205
WA4025	8	238238	247251	369377	281285	304375	166166	178178	269269	267267	217217	211211
WA4027	8	238238	251251	377377	281297	299383	158166	174178	269269	267267	217217	205217
WA4028	8	238238	247251	369377	273273	383395	158166	154194	269269	267267	217217	217217
WA4029	8	238238	251251	377377	269293	375375	166166	162182	269269	267267	217217	205211
WA4030	8	238238	251251	377377	285321	304304	158166	178194	269269	267267	217217	211219
WA4032	8	238238	251251	377377	285293	304379	166166	178194	269269	267267	211217	205211
ER0003	9	238238	251251	369369	285329	000000	166166	000000	000000	000000	211215	215215
ER0007	9	238238	251251	369369	301309	359399	166166	000000	000000	000000	211215	205215
ER0022	9	238238	251251	369369	285301	304359	158166	000000	000000	000000	211215	205215
KE4501	10	226238	247251	369369	297313	304304	166166	186194	269269	267267	211217	205205
KE4502	10	226238	251251	369377	285285	375399	158166	182182	269269	267267	215217	205211
KE4503	10	238238	247251	369369	297313	304304	158166	182202	269269	267267	217217	211217
KE4504	10	238238	251251	369369	293297	304383	166166	186198	269269	267267	211217	205217
KE4505	10	226226	251251	369369	305305	304304	166166	182198	269269	267267	211217	211211
KE4506	10	226238	251251	369377	301317	304304	158166	174182	269269	267267	203217	211217
KE4507	10	226238	251251	369369	305313	308387	166166	000000	269269	000000	215217	217217
KE4508	10	226238	251251	369377	289309	375375	166166	174178	269269	267267	211211	211217
KE4509	10	226238	251251	377377	297301	304304	158166	000000	269269	000000	217217	211217
KE4510	10	238238	251251	369377	289301	308387	166166	000000	269269	000000	211217	205211
KE4511	10	226226	251251	377377	285317	304304	166166	162182	269269	000000	213217	205211
KE4512	10	226238	251251	369369	289293	304391	158166	000000	269269	267267	217217	211211
KE4513	10	238238	247251	369377	281301	304308	166166	170174	269269	267267	213217	211211
KE4514	10	226226	251251	369377	297305	375387	158166	170182	269269	267267	211217	217217
KE4515	10	226238	251251	369377	281285	359411	158166	174186	269269	267267	217217	205217
KE4516	10	238238	251251	377377	293301	304304	158158	170186	269269	267270	211215	205217
KE4517	10	226238	251251	369377	285293	311395	158166	182186	269269	267267	217219	217217

ID	Pop	LAF13	LAF11	EMX4	LaT05	LaT06	LAF10	LAF12	EMX5	EMX3	LAF29	LAF37
KE4518	10	226226	251251	377377	301309	308383	000000	186186	269269	267267	211213	211217
KE4519	10	226226	251251	369377	289293	304304	166166	000000	269269	000000	203215	217217
KE4537	10	226226	251251	369377	305305	375375	166166	000000	269269	267267	215217	205211
KE4539	10	226238	251251	369377	301301	304391	166166	000000	269269	267267	211217	211211
KE4540	10	238238	251251	369377	293301	299299	166166	000000	269269	267267	217217	205205
KE4546	10	238238	251251	369377	289301	304407	158166	178178	269269	000000	215217	211211
KE4548	10	226238	251251	377377	289301	304308	158166	000000	269269	000000	211217	211217
KE4549	10	226226	251251	369369	301309	304304	158166	178182	269269	267267	215217	211217
KE4550	10	226226	251251	369377	293305	304304	166166	178202	269269	267267	211217	205205
KE4601	10	226238	251251	377377	289305	304304	166166	182182	269269	267267	211215	217217
KE4602	10	238238	251251	377377	301313	391395	158158	170178	269269	267267	217217	205217
KE4603	10	226238	251251	369369	289289	304304	166166	182194	269269	267267	207217	205217
KE4604	10	238238	251251	369369	277305	299411	158166	000000	269269	267267	217217	211217
KE4606	10	226226	247251	369369	281297	304304	158166	174190	269269	000000	211217	205211
KE4609	10	226238	251251	369377	297313	304304	158158	186186	269269	000000	217217	205217
KE4610	10	226226	251251	369377	281317	304375	166166	170186	269269	267267	215217	211217
KE4611	10	226238	251251	369369	301301	375391	166166	174194	269269	267267	203211	205217
KE4612	10	226226	251251	369377	301301	304359	166166	178202	269269	267267	211217	205211
KE4613	10	226238	251251	377377	305305	304375	166166	182190	269269	267267	211217	217219
KE4614	10	226238	251251	369377	293297	304395	166166	174194	269269	267267	211217	205211
KE4615	10	238238	251251	377377	293297	304395	166166	178194	269269	267267	203211	205217
KE4616	10	226238	247251	369377	297305	304387	158166	000000	269269	000000	217217	205205
KE4617	10	226238	251251	369377	281285	375399	158166	190190	269269	267267	217217	205211
KE4618	10	226238	251251	369369	301301	304407	166166	170174	269269	000000	211217	205211
KE4619	10	238238	247251	369369	297305	299375	158166	174178	269269	267267	211217	205217
KE4620	10	226238	251251	369377	281297	299308	166166	182186	269269	000000	203217	205211
KE4621	10	226238	251251	369377	281301	375391	166166	174178	269269	267267	217219	205211
KE4622	10	226238	251251	369377	289297	308383	158166	190190	269269	000000	215215	211217
KE4623	10	238238	251251	369377	281301	304304	158166	178210	269269	267267	215219	211211
KE4624	10	226238	251251	369377	297305	304363	158166	162178	269269	267267	215217	205217
KE4625	10	238238	251251	377377	277309	304395	166166	178182	269269	267267	217217	205211
KE4626	10	226238	251251	369377	309309	299304	158166	166186	269269	267267	213219	217217
KE4627	10	226238	251251	369377	293297	304395	166166	178206	269269	267267	217217	205205
KE4628	10	238238	251251	377377	301313	304308	158166	178182	269269	000000	203211	205205
KE4629	10	226238	251251	377377	305305	304379	158166	190202	269269	267267	203217	205205
KE4630	10	226238	251251	369377	301301	299304	158166	162174	269269	267267	213219	211211
KE4721	10	226226	251251	377377	285297	308308	158166	166178	269269	267267	217217	205217
KE4722	10	226238	251251	369377	293305	304387	158166	186198	269269	000000	209217	211211
KE4723	10	226238	251251	369377	285305	299391	158166	182182	269269	267267	217217	205211
KE4725	10	226238	251251	369369	297309	304383	000000	198202	269269	267267	217217	211219
KE4726	10	226226	251251	369369	277297	299375	158166	170182	269269	267270	209215	205213
KE4727	10	226226	251251	377377	297305	304383	158166	000000	269269	000000	213217	211217
KE4728	10	226226	251251	369369	293297	304304	166166	162174	269269	267267	217219	211217
KE4729	10	226238	251251	369377	285305	308383	158166	166186	269269	267267	217217	211217
MK4543	11	238238	247251	369377	289301	304304	166166	000000	269269	267267	215217	211211
MK4544	11	238302	251251	369377	301301	000000	158166	000000	269269	267267	217219	211217
MK4545	11	226238	247251	369377	281309	304375	158166	186198	269269	267267	203215	205211
AB4521	12	226238	251251	369377	301317	308311	166166	162198	269269	267267	217217	211219
AB4522	12	238238	251251	369369	305309	304359	158166	170186	269269	267267	217217	205205

ID	Pop	LAF13	LAF11	EMX4	LaT05	LaT06	LAF10	LAF12	EMX5	EMX3	LAF29	LAF37
AB4523	12	238238	251251	377377	305313	304308	166166	162178	269269	267267	217219	217217
AB4524	12	226238	251251	377377	297313	304308	166166	182182	269269	267267	215217	205205
AB4525	12	238238	251251	369377	285313	304311	158158	178182	269269	267267	215217	205205
AB4527	12	238238	247251	377377	293313	359387	166166	194198	269269	267267	215217	205211
AB4528	12	238238	247251	369369	285289	304379	166166	166194	269269	267267	211217	211211
AB4529	12	226238	251251	377377	297301	299311	158166	170190	269269	267267	217219	205211
AB4530	12	226226	251251	369377	309313	375415	158166	182182	269269	267267	203217	211217
AB4531	12	226238	251251	369377	297317	308359	166166	186222	269269	267267	211215	217219
AB4532	12	226238	251251	377377	281293	308359	158158	162194	269269	267267	213217	205205
AB4533	12	226238	251251	369377	277293	375375	158166	170178	269269	267267	213217	211217
AB4534	12	238238	247251	369377	301309	311375	158158	178182	269269	267267	211217	205211
AB4535	12	238238	247251	377377	265285	304311	158166	166182	269269	267267	211217	205211
AB4541	12	238238	251251	369369	293293	379395	158166	194194	269269	267267	203211	211211
AB4542	12	226238	247251	369377	297305	304304	158158	000000	269269	267267	217217	205211
AM0001	13	226238	247251	377377	281297	304359	158166	182182	269269	267267	217219	205217
AM0003	13	230238	251251	369377	301305	304359	166166	190202	269269	267267	215217	205217
AM0006	13	226226	251251	369377	281289	304367	166166	170182	269269	267267	215217	205211
AM0007	13	238238	247251	377377	281289	291311	166166	178186	269269	267267	211217	205211
AM0008	13	226238	251251	369377	301305	304375	158166	170178	269269	267267	203215	205211
AM0011	13	238238	247251	369377	297301	375383	166166	162190	269269	267267	217217	205205
AM0012	13	238238	251251	369377	297317	304304	158166	182182	269269	267270	211217	205205
AM0014	13	226238	251251	377377	297305	367387	166166	186194	269269	267270	211217	205205
AM0015	13	226230	251251	369377	301309	304383	166166	170206	269269	267267	211217	211211
AM0016	13	226238	251251	369377	297317	304304	166166	178178	269269	267267	211217	205211
AM0018	13	230238	247251	369377	277305	308359	158166	000000	269269	267267	219219	211211
AM0019	13	226226	251251	369377	277305	308308	158158	182182	269269	267267	215219	211217
AM0020	13	226238	251251	377377	305305	304387	158166	178186	269269	267270	211219	205211
AM0021	13	226238	247251	377377	277289	304311	158166	170178	269269	267267	209217	205217
AM0023	13	226238	247251	369377	297329	311311	158158	178182	269269	267267	215215	205217
AM0024	13	230238	251251	377377	301329	291399	158166	190190	269269	267267	211217	205211
AM0025	13	226226	251251	369377	293297	304304	166166	182186	269269	267267	211217	205215
AM0026	13	226230	251251	369377	297297	291304	166166	186186	269269	267267	211217	205205
AM0027	13	238238	251251	369377	305313	304304	166166	182182	269269	267267	209211	205205
AM0028	13	226238	251251	369377	301305	291304	166166	170194	269269	267267	217217	205205
AM0029	13	226238	251251	369377	297305	304311	166166	186190	269269	267267	211217	205213
AM0030	13	226226	247251	369377	297301	304379	158166	178182	269269	267267	211217	205211
AM0031	13	238238	251251	377377	301305	304387	158166	174178	269269	267267	217217	205211
AM0032	13	238238	251251	369377	281301	304311	158166	178182	269269	267267	203215	205205
AM0033	13	238238	247251	369369	293297	311359	158166	174182	269269	267267	217219	205205
AM0034	13	238238	247251	377377	301309	304308	166166	174182	269269	267267	215219	211211
AM0035	13	000000	251251	369377	281305	304304	166166	000000	269269	267267	203217	205217
AM0036	13	226238	247251	377377	297309	304359	158166	182182	269269	267267	215217	205211
AM0037	13	226238	247251	369377	301301	304308	166166	170170	269269	267267	217219	211217
AM0038	13	226226	251251	377377	285305	304383	166166	166178	269269	267267	203215	205205
AM4551	13	226238	251251	377377	289305	291311	162166	166186	269269	267267	217217	205215
AM4554	13	238238	247251	369377	301305	304375	158166	186194	269269	267267	215219	205211
AM4555	13	226226	251251	377377	301313	304304	158166	182190	269269	267267	217219	211211
AM4557	13	226226	000000	369369	285293	304304	158166	182190	269269	267267	211219	205205
AM4558	13	226238	251251	369377	285305	304387	158158	186190	269269	267270	211219	205211

ID	Pop	LAF13	LAF11	EMX4	LaT05	LaT06	LAF10	LAF12	EMX5	EMX3	LAF29	LAF37
AM4559	13	230238	251251	377377	285305	304304	158158	174182	269269	267267	211215	205205
AM4561	13	226226	251251	369377	281293	304304	166166	174182	269269	267267	213215	205211
AM4563	13	226238	251251	377377	281301	304308	158158	190202	269269	267267	217217	211217
AM4564	13	238238	247251	377377	293321	000000	158166	170190	269269	267267	215215	205211
AM4565	13	238238	251251	369377	281313	308391	158158	182190	269269	267267	211217	211211
AM4566	13	226238	251251	369377	277293	304383	166166	174182	269269	267267	213215	205205
AM4567	13	238238	247251	369377	277301	304383	166166	174174	269269	267267	215215	205205
AM4568	13	226238	247251	369377	293305	311311	158162	178186	269269	267267	217217	205215
AM4569	13	238238	247247	369369	293301	311367	158166	178194	269269	267267	215217	205211
AM4570	13	230238	251251	369377	285305	304304	158166	178182	269269	267267	211217	205205
AM4571	13	238238	251251	369377	305329	304304	158166	182190	269269	267267	217217	205205
AM4573	13	226238	247251	377377	305305	304311	166166	190202	269269	267267	215217	205205
AM4574	13	226238	247251	377377	305313	304311	166166	190190	269269	267267	217217	205205
AM4575	13	226238	251251	369377	289301	304387	158166	158162	269269	267267	217217	205205
AM4576	13	238238	251251	369377	301317	304304	158166	158178	269269	267267	211217	205217
AM4577	13	238238	251251	377377	281305	311359	158166	178182	269269	267267	215215	205217
AM4578	13	226238	251251	377377	305317	311371	158158	178182	269269	267267	215217	205205
AM4579	13	238238	247251	377377	281301	304391	166166	182182	269269	267267	217217	205217
AM4580	13	226238	247251	369377	281285	304311	158166	182190	269269	267267	217219	217217
AM4581	13	226238	251251	369369	293293	304304	158166	174182	269269	267267	211217	205211
AM4582	13	238238	251251	369369	281293	304367	166166	174182	269269	267267	217219	205205
AM4583	13	226238	247251	377377	277289	304311	158166	170178	269269	267267	209217	205217
AM4584	13	226226	251251	369377	301301	304387	158158	170178	269269	267267	211215	217217
AM4585	13	238238	251251	369377	285301	304359	158166	178182	269269	267267	215215	205217
AM4587	13	226238	251251	369377	285313	304391	158158	178182	269269	267267	211217	211217
SE2051	14	226226	251251	369369	281297	304304	166166	170170	269269	267267	213215	211217
SE2098	14	226238	251251	369377	289317	304375	158158	170182	269269	267267	217217	205209
SE2100	14	238238	251251	377377	277301	299299	158162	170170	269269	267267	217217	205217
SE2101	14	226238	251251	369377	289297	308407	166166	186194	269269	267267	211217	205217
SE2102	14	238238	251251	377377	289297	304304	166166	178182	269269	267267	211219	217217
SE2103	14	226238	247251	377377	293301	304407	158166	178190	269269	267267	215217	217217
SE2104	14	226226	251251	369369	277301	304304	166166	162186	269269	267267	217217	205211
SE2105	14	230238	251251	369369	281305	359375	166166	178182	269269	267267	215217	205211
SE2106	14	000000	251251	369377	289297	304383	158166	178186	269269	267267	217217	205211
SE2107	14	226226	251251	369377	301309	304304	158166	166178	269269	267267	211211	205217
SE2109	14	226238	251251	369369	277309	304387	158166	166190	269269	267267	211211	217217
SE2156	14	000000	251251	369369	289317	304304	158158	170182	269269	267267	217217	205209
SE2157	14	226238	247251	369377	297305	304387	158158	182202	269269	267267	217219	205211
SE2161	14	226226	251251	369377	289301	304375	158166	162182	269269	267267	215217	217217
SE2162	14	226238	251251	377377	305313	304304	158166	178194	269269	267267	217219	211211
SE2163	14	226226	251251	369377	289305	304304	158166	194206	269269	267267	211215	211217
SE2164	14	226226	251251	369377	293293	304383	166166	182190	269269	267270	211211	205211
SE2165	14	226238	251251	369377	305313	304304	166166	178178	269269	267267	217217	211211
NG2178	15	226238	251251	369369	297313	383395	166166	182194	269269	267267	217217	205211
NG2179	15	226226	247251	369369	285293	304304	158166	174182	269269	267267	215217	217217
NG2180	15	226226	251251	377377	297309	304304	166166	170178	269269	267267	211215	211213
NG2181	15	238238	251251	369377	289309	308308	158166	178194	269269	267267	211217	205205
NG2182	15	226238	251251	369369	297301	308407	166166	182186	269269	267267	217217	205205
NG2191	15	226238	251251	369377	301305	304304	158158	178182	269269	267267	215215	211211

ID	Pop	LAF13	LAF11	EMX4	LaT05	LaT06	LAF10	LAF12	EMX5	EMX3	LAF29	LAF37
NG2192	15	226238	251251	369377	305317	304379	158158	174174	269269	267267	211217	211215
NG2193	15	226238	251251	369377	301305	304383	166166	186194	269269	267267	211219	205211
NG2194	15	226238	251251	369377	297313	304375	166166	182202	269269	267267	217217	217219
NG2207	15	238238	251251	377377	277313	304387	158166	170198	269269	267267	215217	205205
NG2208	15	238238	247251	369369	301301	304304	166166	178190	269269	267267	215217	211217
NG2214	15	226238	247251	369377	301309	304304	166166	178186	269269	267267	211211	205217
NG2215	15	226238	247251	369377	289289	304407	158166	174174	269269	267267	215217	211211
NG2229	15	238238	247251	369369	289289	299304	166166	174194	269269	267270	211215	205211
TA1143	16	226238	251251	369377	293309	304375	166166	166190	269269	267267	215217	205211
TA1144	16	238238	251251	369377	289297	304304	158166	182186	269269	267270	217217	211217
TA1145	16	226238	247251	369377	301301	299383	158166	174174	269269	267267	211211	211217
TA1431	16	226226	247251	369377	289305	304375	000000	190202	269269	267270	213217	205211
TA1432	16	226238	251251	369377	305309	304304	166166	186190	269269	267267	215215	205217
TA1436	16	226226	251251	369377	285309	299383	166166	174194	269269	267267	211217	205205
TA1438	16	226238	247251	369377	289305	375403	166166	190202	269269	267267	215217	217217
TA1439	16	226226	251251	369377	297305	308311	158158	186190	269269	267267	217219	205211
TA1440	16	226238	251251	369369	305313	304304	166166	178186	269269	267267	211217	205205
TA1441	16	226226	251251	369377	309313	375383	166166	170190	269269	267267	211217	211217
TA1443	16	226238	251251	369377	293309	299375	166166	170190	269269	267267	211215	217217
TA1446	16	226238	251251	369369	289313	375395	158166	174178	269269	267267	217217	205205
TA1447	16	226226	251251	369369	301345	304359	158166	174194	269269	267267	215217	205205
TA1449	16	226226	251251	369377	293313	304375	166166	174178	269269	267267	215215	205205
TA1450	16	226226	251251	369377	293301	304383	166166	174178	269269	267267	211211	205211
TA1452	16	238238	251251	369377	297301	304375	166166	170174	269269	267267	217217	205211
TA1454	16	226226	251251	369377	277309	304308	166166	178186	269269	267267	203217	205211
TA1455	16	226238	247251	377377	301313	304395	158166	000000	000000	000000	211219	211215
TA1458	16	226238	251251	369377	289313	304304	166166	178186	269269	267267	211217	205205
TA1459	16	238238	247251	369369	297305	304375	158166	174190	269269	267267	217217	211217
TA1460	16	226226	251251	369377	281305	308375	166166	186190	269269	267267	211215	205211
TA1466	16	226238	251251	377377	289309	304375	158166	178190	269269	267267	217219	205205
SW0901	17	226238	247251	369369	277289	000000	166166	182186	269269	267267	211219	205205
SW0902	17	226238	251251	369377	297301	304379	158166	174194	269269	267267	217219	205211
SW0903	17	226238	251251	377377	281301	304304	166166	170170	269269	267267	217217	211217
SW0904	17	238238	251251	369369	285289	304391	166166	178182	269269	267267	211211	211217
SW0905	17	238238	247251	369369	281305	304304	166166	182194	269269	267267	219221	217217
SW0906	17	226238	247251	369377	297309	304391	158166	178194	269269	267267	211217	211217
SW0907	17	226238	251251	369377	289309	375383	158166	174190	269269	267267	211217	205211
SW0908	17	238238	251251	369369	297321	304375	158166	190194	269269	267267	217219	205211
SW0909	17	226238	251251	369377	273301	304304	166166	174178	269269	267267	215217	203211
SW0910	17	238238	000000	369369	273305	304304	166166	194198	269269	267267	215219	205217
SW0911	17	226226	247251	369377	285289	304367	158166	174186	269269	267267	211215	211211
SW0912	17	238238	251251	369377	281289	304304	166166	182194	269269	267267	219219	000000
SW0913	17	226226	251251	377377	277305	304304	166166	166178	269269	267267	211219	205211
SW0914	17	230238	247251	369369	297305	375383	166166	182186	269269	267267	217217	203203
SW0915	17	226238	251251	369369	277313	308311	166166	166174	269269	267270	217217	203203
SW0916	17	226238	251251	369377	301341	304311	166166	182182	269269	267267	211217	211217
SW0917	17	226226	251251	369369	305313	304304	166166	182190	269269	267267	211217	211219
SW0918	17	226226	251251	377377	277309	304304	166166	174190	269269	267267	217219	205211
ZZ0145	18	226238	251251	369377	293301	311383	166166	186186	269269	267267	211219	205217

ID	Pop	LAF13	LAF11	EMX4	LaT05	LaT06	LAF10	LAF12	EMX5	EMX3	LAF29	LAF37
ZZ0147	18	226238	251251	369377	277313	304304	166166	162190	269269	267270	217217	205211
ZZ0148	18	226238	251251	377377	273289	304304	166166	182186	269269	267267	217219	205211
ZZ0149	18	238238	247251	369377	281305	304304	166166	182186	269269	267267	211217	205211
ZZ0155	18	226238	247251	000000	289305	304308	166166	178194	269269	267267	211215	205211
ZZ0157	18	226226	251251	369369	277285	304308	166166	174194	269269	000000	217217	205217
ZZ0159	18	238238	251251	000000	273285	304304	166166	162186	269269	267267	211217	213217
ZZ0160	18	226238	247251	000000	293293	304383	166166	178190	269269	267267	211217	217217
ZZ0161	18	226238	251251	000000	289305	304304	166166	182190	269269	267267	215215	205211
ZZ0162	18	226238	247251	000000	285293	304304	166166	182194	269269	267267	217217	211217
ZZ0163	18	238238	247251	000000	277289	304304	166166	182190	269269	267267	215217	217217
ZZ0164	18	226238	251251	000000	321321	304304	158158	186186	269269	267267	000000	205205
HW0049	19	226238	251251	377377	301321	304407	166166	000000	269269	267267	215219	211211
HW0057	19	238238	251251	377377	305321	355375	166166	174178	269269	267270	217217	205211
HW0059	19	238238	251251	369369	297305	304355	000000	194198	269269	267267	211217	205211
HW0060	19	226238	251251	377377	289313	304304	158158	186198	000000	267267	215217	205211
HW0061	19	226238	251251	369377	293301	379387	166166	186186	000000	267270	215217	205211
HW0062	19	226238	251251	369377	277285	304375	166166	194194	269269	267267	211211	205217
HW0063	19	226238	251251	369377	301313	304304	158166	174186	269269	267267	211211	211217
HW0067	19	226226	251251	369377	273301	304308	158166	182190	000000	267270	211217	205211
HW0068	19	000000	251251	377377	277293	304304	166166	178178	269269	267267	207217	205211
HW0069	19	226238	251251	369377	277317	304383	166166	178186	000000	267267	211211	205205
HW0072	19	238238	251251	369377	289293	304304	166166	190190	000000	267267	211217	205211
HW0076	19	238238	251251	369377	289297	304379	166166	174194	000000	267270	217217	205211
HW0081	19	226226	000000	369377	301301	304308	166166	174182	000000	267267	211217	205211
HW0082	19	226226	251251	369369	277325	304367	158166	186186	000000	267270	000000	205211
HW0083	19	226238	251251	369377	297305	304304	166166	182182	000000	267267	217217	205211
HW0084	19	238238	251251	369369	289297	304304	166166	186186	269269	267270	215219	211217
HW0086	19	226226	251251	369369	301305	304308	158166	166178	269269	267267	207215	205205
HW0088	19	226238	251251	369377	277281	304383	166166	186194	269269	267267	211219	211217
HW0089	19	226238	247251	377377	281289	304304	166166	170186	269269	267267	211217	205217
HW0090	19	226238	251251	369369	277293	379383	158166	186190	269269	267267	217219	205211
HW0091	19	226226	251251	369377	281305	304304	158166	166178	269269	267267	211217	211211
HW0092	19	000000	251251	369377	289301	304304	166166	186198	269269	267267	211217	205205
HW0097	19	226230	247251	369369	273305	304304	166166	182182	269269	267270	215217	217217
HW0099	19	226238	247251	369377	289301	304304	166166	162186	269269	267267	209211	205211
HW0101	19	226226	251251	369377	281301	304304	158166	178190	269269	267267	207217	205211
HW0102	19	226238	251251	369369	293313	304367	166166	182186	269269	267267	211217	205217
HW0103	19	226226	251251	377377	293305	304375	166166	182194	269269	267267	211217	205217
HW0112	19	226238	251251	369369	281297	308375	158166	182186	269269	267267	211211	211217
HW0113	19	226238	251251	369369	273277	304304	166166	162178	269269	267267	217217	211211
HW0116	19	238238	000000	369369	277301	304304	158166	182182	269269	267267	211217	205217
HW0117	19	226238	251251	377377	273313	304304	166166	182190	269269	267267	211219	217219
HW0118	19	000000	251251	369377	277321	304304	166166	186186	269269	267270	211211	211211
HW0119	19	238238	251251	369377	273317	304387	166166	182186	269269	267267	215215	211217
HW0120	19	238238	251251	369377	293325	304379	166166	178182	269269	267267	217217	205217
HW0122	19	226238	247247	369377	305313	367391	166166	182194	269269	267270	211211	205211
HW0123	19	226238	251251	369369	297313	304367	158166	182186	269269	267267	217217	205211
HW0124	19	226226	251251	369369	293325	304304	158166	186194	269269	267267	217217	211211
HW0138	19	238238	251251	369369	285305	308311	166166	182186	269269	267270	209217	205211

ID	Pop	LAF13	LAF11	EMX4	LaT05	LaT06	LAF10	LAF12	EMX5	EMX3	LAF29	LAF37
HW0151	19	226226	251251	369369	297305	304311	166166	174186	269269	267270	207217	205205
CH0878	20	226238	251251	369369	289301	304311	166166	166194	269269	267267	217219	205211
CH0882	20	226238	251251	369377	277305	304304	158166	186186	269269	267267	211217	205211
CH0884	20	226226	251251	377377	277317	308395	166166	166182	269269	267267	215217	211217
CH0885	20	000000	247251	369369	297305	304304	166166	186186	269269	267270	215217	205205
CH0890	20	226226	251251	369377	273273	304304	166166	162178	269269	267267	217217	205205
CH0891	20	226238	251251	369369	273313	304308	166166	178190	269269	267267	211217	205217
CH0892	20	226238	251251	377377	277281	304304	166166	182186	269269	267270	211215	205211
CH0894	20	238238	251251	369377	281293	304383	158166	166182	269269	267267	215219	205205
CH0895	20	238238	247251	369369	277313	308399	158166	174178	269269	267270	215217	205211
CH0897	20	226238	251251	369377	281289	304391	000000	174182	269269	267270	217219	205205
CH0898	20	000000	247251	369369	289305	304304	166166	190190	269269	267267	211217	205211
CH0899	20	000000	251251	369369	281293	304311	158158	182186	269269	267267	211217	205211
CH0906	20	226238	251251	369377	285325	304304	166166	174182	269269	267270	211217	211219
CH0908	20	226238	251251	369377	297325	304403	166166	178186	269269	267267	211219	205211
CH0931	20	238238	251251	369377	305321	304371	158166	178186	269269	267267	211219	205211
CH0934	20	226238	247251	369369	289301	304304	166166	174194	269269	267267	217217	205205
CH0935	20	226226	251251	369377	281285	000000	158166	190190	269269	267267	211217	205211
CH0953	20	226238	247251	369369	285297	304304	166166	000000	269269	267267	215217	205205
CH0958	20	226238	247251	369377	277289	304379	158166	182186	269269	267267	211215	205217
CH0959	20	226238	251251	369377	273285	375387	158166	186190	269269	267267	217217	205217
CH0960	20	226226	251251	369369	293297	304359	166166	166198	269269	267267	211217	205219
CH0962	20	226238	251251	377377	269293	304387	166166	166190	269269	267270	217219	000000
CH0963	20	238238	251251	369369	281317	304304	166166	182182	269269	267267	211219	205211
CH0964	20	226226	251251	377377	289297	304304	166166	178182	269269	267267	215217	205211
CH0965	20	226238	251251	369369	289293	295359	158166	178194	269269	267270	217217	205211
CH0967	20	226238	251251	369377	285285	304304	166166	182190	269269	267267	211219	205211
CH0968	20	226238	251251	369377	285297	304304	158166	000000	269269	267267	211219	205205
CH0969	20	226238	251251	369377	305313	304383	166166	186190	269269	267267	217219	205211
CH0970	20	226238	251251	369377	289293	304391	166166	182186	269269	267267	209215	205211
SA0972	21	238238	251251	369369	281285	304311	166166	182186	269269	267267	217217	205213
SA0976	21	238238	247251	369377	289305	304304	166166	178186	269269	267267	211211	211217
SA0989	21	226238	251251	369377	305305	304304	158166	000000	269269	267267	217217	217217
SA0990	21	226238	251251	369377	277305	304308	158166	182186	269269	267270	217217	205205
SA0991	21	226226	251251	377377	301305	304304	166166	000000	269269	000000	209211	205205
SA0993	21	226238	247251	377377	301305	304311	166166	170182	269269	267270	211217	205213
SA0994	21	226226	251251	377377	289293	304311	166166	182182	269269	267267	211211	205205
SA0995	21	226238	251251	369377	277301	304383	158166	182190	269269	264267	217217	205217
SA0996	21	226226	251251	369377	309317	304304	166166	190190	269269	267267	211215	205217
SA1002	21	226238	251251	369369	313321	304304	166166	174186	269269	267267	209211	205217
SA1003	21	226238	251251	369377	297317	304311	166166	170182	269269	267267	211219	205211
SA1004	21	238238	243251	369369	285289	375383	166166	000000	269269	267267	217217	205205
SA1005	21	226226	251251	369377	281293	304311	166166	182182	269269	267267	211219	205217
SA1006	21	226238	251251	369377	285305	304304	158158	174186	269269	267267	211217	205217
SA1008	21	226238	251251	369369	309309	304399	166166	170178	269269	267267	215217	205211
SA1009	21	226238	247251	377377	281305	304304	158166	182190	269269	267267	211217	205205
SA1010	21	226238	251251	369377	289297	304304	166166	178182	269269	267267	211217	205217
SA1012	21	238238	251251	377377	305317	304383	166166	166186	269269	000000	217217	205217
SA1015	21	000000	251251	369369	293297	304304	166166	000000	000000	000000	217217	205205

ID	Pop	LAF13	LAF11	EMX4	LaT05	LaT06	LAF10	LAF12	EMX5	EMX3	LAF29	LAF37
SA1025	21	226238	247251	369369	277297	304311	166166	174182	269269	267270	211219	211213
SA1027	21	226226	251251	369369	309325	304304	166166	000000	269269	267267	217217	211211
SA1028	21	238238	251251	369377	305313	304304	166166	178190	269269	267267	211217	205205
SA1029	21	238238	247251	369369	277305	304304	166166	182190	269269	267267	211217	217219
SA1032	21	226238	247251	369377	297309	304304	158166	000000	269269	267270	207211	205211
SA1033	21	226226	251251	369369	305309	304375	158166	178186	269269	267267	211219	211211
MA0802	22	226238	251251	369369	289309	304308	158166	178186	269269	270270	211211	205205
MA0803	22	226238	247251	369377	297305	363391	166166	174186	269269	267267	217217	211217
MA0804	22	226238	251251	369369	281313	308375	158166	178194	269269	267267	211215	205205
MA0806	22	226238	251251	369377	277317	304308	158158	178178	269269	267267	215215	205211
MA0807	22	226226	243251	369377	301301	304371	158158	178194	269269	267270	205217	205211
MA0808	22	226238	251251	369369	281297	308375	158166	178194	269269	267267	211215	205217
MA0810	22	238238	251251	369369	281313	304371	158166	186186	269269	267270	217219	205211
MA0811	22	226238	251251	369377	313321	304304	158166	178178	269269	267270	213219	205217
MA0812	22	238238	251251	369369	281301	308308	166166	170186	269269	267267	211215	211217
MA0813	22	226238	251251	369369	289313	308308	166166	174182	269269	267270	217217	205217
MA0814	22	226238	243247	369369	297301	375383	166166	186194	269269	267267	211217	205205
MA0815	22	226226	243251	369369	281301	308308	158166	174182	269269	267267	213215	205215
MA0816	22	226238	247251	377377	281305	308308	166166	182186	269269	267267	211219	211217
MA0817	22	226238	251251	377377	305317	308308	166166	182182	269269	267267	211215	205211
MA0818	22	226238	251251	369377	301309	308403	158166	178194	269269	267270	217219	205211
MA0819	22	238238	243251	369369	281289	304383	166166	178186	269269	267270	217217	205211
MA0820	22	226238	251251	369369	281313	308403	158166	182182	269269	267270	217217	205217
MA0821	22	238238	251251	369377	281281	304308	166166	190194	269269	267267	211217	205205
MA0822	22	226226	247251	369377	305313	304308	158166	190194	269269	267270	217217	217217
MA0823	22	226226	247251	369377	297297	304304	166166	190194	269269	267270	217217	203217
MA0824	22	226238	247251	369377	297313	304367	166166	178178	269269	267267	215217	217217
KR0005	23	238238	251251	377377	297301	304411	166166	186206	269269	267267	213217	211211
KR0006	23	226238	251251	000000	297321	304304	158166	162186	269269	267267	211217	211211
KR0007	23	226226	251251	000000	281297	304311	166166	174178	269269	267267	215217	205205
KR0008	23	226238	247251	000000	297309	304304	158166	182186	269269	267267	215217	205205
KR0009	23	226238	251251	000000	297297	304308	158166	000000	269269	267267	213217	205205
KR0011	23	226226	251251	000000	289313	304308	158166	170174	269269	267267	217217	217217
KR0013	23	238238	247251	000000	305325	304415	166166	178182	269269	267267	217219	205217
KR0014	23	238238	251251	000000	297297	304379	158166	186194	269269	267267	211217	211217
KR0015	23	226238	251251	369377	297317	304379	166166	000000	269269	000000	211219	205211
KR0017	23	226238	247251	000000	321325	304304	000000	178182	269269	267267	217219	205217
KR0018	23	238238	251251	369369	297325	304304	166166	174186	269269	267267	217217	211217
KR0021	23	226238	251251	377377	281305	308403	166166	000000	269269	267267	217217	205211
KR0024	23	226238	251251	369377	293297	304395	158166	182198	269269	000000	211215	205217
KR0027	23	238238	251251	369369	301301	304304	158166	170174	269269	267267	211211	211217
KR0028	23	226238	251251	369377	297301	000000	158166	174174	269269	267267	211217	205217
KR0032	23	238238	251251	369377	297305	304311	158166	170170	269269	267267	217219	205211
KR0034	23	238238	247251	369377	297301	304379	166166	162182	269269	267267	213217	205211
KR0035	23	238238	247251	369377	301301	304311	166166	174182	269269	267267	211219	211217
KR0041	23	226238	247251	369369	309309	304387	166166	186202	269269	267267	211219	205211
KR0043	23	238238	247251	369377	297321	304304	166166	174198	269269	267267	215215	205211
KR0049	23	226238	247251	369369	297309	379387	158166	170206	269269	267267	217219	205205
KR0051	23	226238	251251	000000	293297	304304	158166	178186	269269	267270	207211	211211

ID	Pop	LAF13	LAF11	EMX4	LaT05	LaT06	LAF10	LAF12	EMX5	EMX3	LAF29	LAF37
KR0057	23	238238	247251	369369	301301	308308	166166	182186	269269	267267	215217	205217
KR0061	23	238238	251251	369377	297325	308379	166166	000000	269269	267267	217219	205211
KR0068	23	238238	251251	369377	285297	304304	158158	174186	269269	267267	211211	205211
KR0071	23	238238	251251	369377	285293	304304	166166	000000	269269	267267	211211	205211
KR0074	23	238238	247251	377377	297317	304308	166166	194198	269269	267267	211213	205209
KR0078	23	226238	251251	377377	297325	304304	166166	000000	269269	267267	211217	205217
KR0090	23	226238	247251	369369	297297	308308	158166	162162	269269	267267	213217	205205
KR0094	23	238238	247251	377377	277301	308308	158166	000000	269269	267267	211217	205205
KR0096	23	238238	251251	369369	301313	304308	166166	000000	269269	267270	207217	205211
KR0099	23	238238	251251	377377	313321	304304	166166	000000	269269	267267	213213	205211
KR0107	23	226238	251251	369377	293305	308379	158166	000000	269269	000000	213217	211217
KR0108	23	238238	251251	369377	297305	304403	166166	186194	269269	267267	211217	211217
KR0109	23	238238	251251	369377	293313	304308	166166	182182	269269	000000	217217	205211
KR0116	23	238238	251251	369377	301309	304308	166166	000000	269269	267267	215217	211217
KR0118	23	238238	251251	369377	293313	304304	166166	000000	269269	267267	215217	205211
KR0123	23	226238	251251	377377	273297	304308	158166	000000	269269	267267	215217	211211
KR0128	23	226238	251251	369377	305325	304308	158166	162186	269269	267267	217219	211211
KR0129	23	238238	251251	369369	277281	304304	166166	000000	269269	000000	211211	205217
KR0130	23	226238	251251	369377	277325	304304	166166	000000	269269	267267	211217	205217
KR0137	23	226238	251251	369377	281301	304308	166166	000000	000000	000000	215215	211211
KR0138	23	226238	251251	369377	297301	308375	158166	170182	269269	267267	215217	211211
NA4651	24	238238	251251	369369	285289	363379	166166	174186	269269	267267	211217	205217
NA4652	24	226226	251251	000000	277289	304304	166166	174182	269269	267270	217217	211217
NA4653	24	238238	247251	369377	293305	304304	166166	190222	269269	267270	211211	211211
NA4655	24	226238	247251	369377	301309	304311	158166	174174	269269	267267	211217	205211
NA4656	24	238238	251251	369377	309309	304375	166166	178186	269269	267267	211217	205211
NA4658	24	226238	251251	000000	293297	363379	158166	174174	269269	267270	205217	205211
NA4659	24	226226	251251	000000	293309	363375	166166	186186	269269	267267	211217	205211
NA4662	24	238238	247251	369377	305309	304375	166166	174186	269269	267267	211217	205211
NA4664	24	226226	251251	000000	305317	383395	166166	178226	269269	267270	211217	211217
NA4665	24	226238	251251	000000	293297	304363	166166	174198	269269	267267	211217	217217
NA4666	24	226226	247251	000000	285309	304379	166166	174182	269269	267267	211211	205217
NA4667	24	226238	251251	000000	285309	304375	158166	190222	269269	267267	211217	205205
NA4668	24	226226	247247	000000	277293	304304	158166	186190	269269	267270	217217	211211
NA4669	24	226238	251251	000000	285297	304304	166166	186186	269269	000000	211217	211217
NA4670	24	238238	251251	000000	309313	383395	166166	170174	269269	267267	217219	205211
NA4671	24	226226	251251	000000	309309	363383	166166	182182	269269	267267	211217	205205
NA4672	24	238238	247251	377377	285297	375379	158166	190190	269269	267267	217217	217217
NA4673	24	226238	247251	000000	293305	363391	166166	186186	269269	267270	207219	205211
NA4674	24	226226	251251	000000	285293	304379	166166	178186	269269	267267	211217	217217
NA4675	24	238238	247251	000000	277297	304304	166166	178178	269269	267267	211217	205211
NA4677	24	238238	251251	000000	305313	304304	166166	000000	269269	267267	211215	205211
NA4678	24	238238	247251	000000	305325	311379	166166	174178	269269	267267	217217	205217
NA4679	24	226238	247251	000000	285293	304363	166166	178190	269269	267267	217217	205211
NA4680	24	226238	251251	000000	301317	304304	158166	174186	269269	267270	217217	205205
NA4685	24	226226	251251	377377	309309	304304	166166	186190	269269	267267	217217	205211
NA4686	24	226238	251251	369369	297309	375395	166166	174190	269269	267270	217217	211217
NA4687	24	238238	251251	369369	297297	304304	166166	186186	269269	267267	217217	211211
NA4689	24	238238	251251	369377	297309	304375	166166	174178	269269	000000	217217	217217

ID	Pop	LAF13	LAF11	EMX4	LaT05	LaT06	LAF10	LAF12	EMX5	EMX3	LAF29	LAF37
NA4690	24	226238	251251	369377	293293	304311	166166	174174	269269	267267	217217	217217
NA4691	24	226226	247251	369377	293293	304367	166166	186186	269269	267267	207211	211217
NA4692	24	226226	247251	377377	277305	379379	166166	174186	269269	267270	205207	211211
NA4695	24	226238	251251	369377	301309	304375	166166	178182	269269	267270	205211	211211
NA4696	24	226238	247251	369377	293297	304308	166166	174190	269269	267267	207217	211217
NA4697	24	226238	247251	369377	277293	379379	166166	174186	269269	267267	211211	205205
NA4698	24	226226	251251	369369	293297	304363	166166	174190	269269	267270	217217	205211
NA4699	24	226238	251251	369377	289305	304304	166166	174186	269269	267267	217217	211211
NA4702	24	226226	247251	369369	309321	379391	166166	182186	269269	267267	211217	211217
NA4703	24	226238	251251	369369	281309	304395	166166	190190	269269	267267	211217	205205
NA4704	24	226226	251251	369369	277285	308363	166166	182190	269269	267270	217217	211217
NA4706	24	226238	251251	369369	289309	304363	166166	174178	269269	267267	211217	205217
NA4707	24	226238	251251	369369	277317	363391	166166	000000	269269	267267	211217	211211
NA4708	24	226238	251251	369377	309309	304304	166166	174186	269269	267270	205211	205211
NA4709	24	238238	247251	369377	281305	304304	166166	190222	269269	270270	217217	205205
NA4710	24	226238	251251	369369	309309	304304	166166	000000	269269	000000	205211	205211
NA4711	24	238238	247251	369369	293309	363367	166166	182190	269269	267270	211219	205211
NA4712	24	238238	247247	369369	293293	383395	158166	182190	269269	267267	211217	211217
NA4713	24	226238	251251	369377	293301	304363	166166	186190	269269	267267	211211	205205
NA4717	24	226238	251251	369369	293297	367375	158166	182190	269269	267267	211217	205211
NA4721	24	226238	247251	369369	301309	304311	158166	000000	269269	267267	211217	205217
NA4722	24	238238	251251	369377	277297	304304	166166	174186	269269	267267	215217	205205

APPENDIX B: DNA ENRICHMENT FROM LOW QUALITY SAMPLES BY CAPTURE HYBRIDIZATION FOR USE IN WILDLIFE CONSERVATION GENETICS

Abstract

Studying the genetics of rare, threatened and endangered species has been greatly improved through the use of non-invasively collected DNA samples (i.e., dung). However the quality of DNA from these samples is low, limiting the genetic markers that can be analyzed. Methods that can increase the utility of low quality DNA sources have been introduced and work through the combination of DNA enrichment by hybridization capture and next generation sequencing. Here we adapt methods to enrich non-invasively collected DNA from dung for nuclear and mitochondrial loci useful in conservation genetics research. Using African elephants as a model, we developed capture probes for enrichment of low quality DNA sources (such as dung and ivory) for entire mitochondrial genomes; X-linked genes *BGN*, *PLP*, and *PHKA2*; Toll-like receptor genes *TLR 3, 7* and *8*; and the major histocompatibility complex gene *DQA*.

Introduction

Despite advances in research technology and management practices countless species of wildlife continue to decline and become threatened with extinction, largely due to human activities like poaching. The research field of conservation genetics aims to reduce the rate and severity of species decline by better understanding the genetics of species of interest and making the relevant information available to management and law enforcement agencies. High quality DNA is required for many of the currently available techniques from which important genetic information can be obtained. Blood and tissue are the most common sources of high quality DNA in mammals, and collecting these types of samples typically requires physical contact and stressful handing of the animal. For species that are rare or cryptic, adequate sampling is difficult and often specialized training, numerous permits, and costs make sufficient sampling unrealistic. Recently, samples that yield lower quality DNA, such as dung, have gained popularity due to the relatively easy collection process, abundance of samples, and lack of direct contact with individuals during sampling. However, DNA isolated from dung can be highly fragmented while containing non-target DNA (e.g., DNA from microbes, plants or prey items) and numerous PCR inhibitors. To overcome these obstacles, techniques have been developed to increase the utility of DNA from dung, such as the development of markers with short PCR amplicons and the inclusion of PCR additives that reduce inhibition.

Here we adapt methods to increase the genetic information that can be obtained from low quality DNA sources for application to conservation genetics research using DNA enrichment methods and next generation sequencing technology. Specific regions of DNA can be targeted for sequencing in samples containing inhibitors and non-specific DNA, by hybridization and immobilization to a biotin-labeled capture probe. Using this technique to enrich and purify DNA from low quality sources, numerous loci can be extracted and efficiently sequenced using next-generation sequencing platforms such as the Illumina MiSeq. For this study we utilize African elephants, as they are an ideal test case with direct conservation implications aided by this research. African elephants are charismatic mega fauna that are at risk of extinction due to poaching. Law enforcement agencies typically intercept shipments of ivory outside of elephant range countries, and as a result the location from which the ivory was poached can be difficult to ascertain. The research proposed here will demonstrate not only a way to increase the amount of

genetic information obtainable from low quality and degraded DNA sources, but will make significant contributions to ongoing research attempting to establish the provenance of confiscated illegal ivory.

Elephants as a Model

There is strong evidence for two extant species of African elephant, the African savanna elephant (*Loxodonta africana*) and the African forest elephant (*Loxodonta cyclotis*) (Brandt *et al.* 2012; Ishida *et al.* 2011b; Roca *et al.* 2005; Roca *et al.* 2001; Rohland *et al.* 2010). African forest elephants occupy tropical rainforests while African savanna elephants are found in savanna bush and lightly forested regions (Grubb *et al.* 2000). Forest and savanna elephants can be distinguished from one another in the field by several features; savanna elephants have large triangular shaped ears, as well as forward and outward curved tusks, whereas forest elephant ears are smaller and rounded, the tusks are thinner and straighter, and the overall body shape is smaller and more compact (Grubb *et al.* 2000).

African elephants were listed as threatened under the Endangered Species Act in 1978 and were awarded further protection in 1989 when they were added to Appendix I of the Convention on International Trade in Endangered Species of Wild Flora and Fauna (CITES), thus prohibiting commercial trade of elephant products. Despite this ban poaching continues to be a major threat to elephant conservation (Barnes 1999; Douglas-Hamilton 1987; Wasser *et al.* 2009) and large quantities of ivory continue to be smuggled illegally (Species Survival Network 2007). Using molecular data to track the origin of illegal ivory in an effort to stem poaching has been a long-term goal in elephant conservation (Comstock *et al.* 2002; Ishida *et al.* 2013; Wasser *et al.* 2007).

Methods have been developed that utilize nuclear short tandem repeats (STRs) (Wasser *et al.* 2009; Wasser *et al.* 2008; Wasser *et al.* 2007; Wasser *et al.* 2004), single nucleotide polymorphisms (SNPs) (Ishida *et al.* 2011a), and mitochondrial DNA (Ishida *et al.* 2013) to establish the origin of seized shipments of illegal ivory. The nuclear STRs and SNPs have the ability to distinguish whether ivory samples originated from savanna or forest elephants with near perfect accuracy, and population origin with some success when compared to reference samples of known provenance. African elephants have a social structure that influences the geographic distribution of nuclear and mitochondrial DNA alleles; males mediate nuclear gene

flow across the landscape as they migrate away from their natal herd, while maternally inherited mitochondrial DNA remains with the herd by non-dispersing females (Ishida *et al.* 2011b; Roca *et al.* 2005). The phylogenetic information that results from the constraint on mitochondrial DNA was examined by Ishida and colleagues (2013), who found that distinctive haplogroups that correspond to geographic regions within the African continent and unique haplotypes that were even more limited in their distribution. This provided for genetic methods for determining the origin of confiscated ivory.

Efforts to develop methods for establishing the source of ivory have faced two major hurdles. (1) DNA isolated from ivory is variable in quality and quantity (Mailand & Wasser 2007). To account for this variation, nuclear and mitochondrial markers are best designed with short PCR amplicons (< 300 bp); however this significantly reduces the amount of genetic information that can be obtained from a single ivory sample. (2) In order to determine the source of ivory, a large database of reference samples is needed. Traditionally, genetic studies utilize high quality DNA sources like blood or tissue; however, obtaining these types of samples can be costly and require specialized training, permits and physical interaction with the animal. Using non-invasively obtained samples like dung eliminates the need for stressful handling of individuals during sample collection and increases the proportion of a population that can be sampled. DNA isolated from dung has its shortcomings: it is highly fragmented, contains a large proportion of non-target DNA, has numerous PCR inhibitors and often only markers that target short regions can be successfully amplified. To circumvent these limitations and increase the utility of DNA from ivory and dung samples, we have adapted methods for enriching DNA by capture hybridization from low quality elephant samples for use with next generation sequencing technology. Studies on African elephants using DNA enrichment will maximize the value of noninvasively collected dung and confiscated ivory samples for answering a variety of research questions involving population origin, structure, and diversity.

Objectives

The main goal of this study was to develop cross-species capture probes capable of enriching for both nuclear and mitochondrial DNA markers useful in answering a wide range of population and conservation genetics questions for rare, threatened or endangered species. DNA enrichment techniques have been applied to improve the efficiency of next generation

sequencing technology. Most studies have used enrichment and next-generation sequencing for ancient DNA, which due to taphonomic processes is degraded much like DNA from dung or ivory, and DNA from modern non-human primates. Mason and colleagues (2011) have demonstrated the utility of cross-species capture hybridization in colugos using DNA from museum skins. They were able to enrich for mitochondrial DNA fragments and sequence the entire mitochondrial genome using capture probes with as much as 13% sequence divergence from their target species. Hybridization capture was also used to sequence full mitochondrial genomes from 5000 year old human skeletal remains (Cui *et al.* 2013), and to sequence 1.5 megabases of chromosome 21, chromosome X, and the complete mitochondrial genome from western chimpanzee blood and feces (Perry *et al.* 2010). These studies relied on commercially available enrichment kits to develop capture probes, which requires the use of a reference genome. Few complete genome sequences are available, thus limiting the use of commercially available enrichment kits. If capture probes developed from related species can be used to enrich for nuclear DNA together with mitochondrial DNA, then the genetic information that can be obtained from low quality DNA sources like dung will be substantially improved.

African elephants serve well as a model and test case for cross-species capture probe development. Elephant genetics have been extensively investigated using loci commonly implemented in the genetic study of other species. In this study we developed capture probes and attempted an efficient enrichment method for: (1) entire mtDNA genomes; (2) nuclear genes, including X-linked genes *BGN*, *PLP* and *PHKA2*, (3) toll-like receptor immune system, including *TLR3*, *TLR7* and *TLR8*; and (4) the major histocompatibility complex gene *DQA*. The amplicons for these loci are large (>300 bp), and as such would typically be difficult to amplify directly from low quality DNA sources. However, using DNA enrichment by capture hybridization and subsequent next generation sequencing these markers could potentially be successfully examined in DNA from dung.

Materials and Methods

Sample Collection and DNA Isolation

Dung was collected from different herds at various locations in Botswana and Guinea to minimize resampling the same individual. Samples were taken from the exterior of fresh dung (~1 to 2 days old at most); as this is the area that is most like to have epithelial cells that have sloughed off from the elephant's large intestine. Collected dung was heated to 72 ^{o}C for at least 30 minutes to comply with requirements from the United States Department of Agriculture – Animal and Plant Health Inspection Service (USDA-APHIS), for treatment of Foot and Mouth Disease (FMD) prior to importation. Once treated, buffer (salt saturated 20% dimethyl sulfoxide solution with 100 mM Tris and 0.25 M EDTA) was added to the samples for storage and shipment to the United States. This buffer has been shown to preserve DNA for up to one year without needing refrigeration (Amos *et al.* 1992; Eggert *et al.* 2008). Samples were stored at room temperature or refrigerated until shipped. Once received, DNA was immediately extracted using the QIAamp DNA Stool Kit (Qiagen Inc., Valencia, CA) following the recommended protocol.

Illumina Library Preparation

Following isolation, to ensure optimal fragment size, the DNA samples were sheared by sonication to 800 bp using the M220 Foucused-ultrasonicator (Covaris, Inc. Woburn, MA) in 130 ul volumes in microTUBE AFA Fiber Snap-Cap tubes. Library preparation was performed using the NEXTflex DNA Sequencing Kit (BIOO Scientific Corp., Austin, TX) with modifications to the manufacturer protocol. Gel-free size selection utilizing AmpureXP beads (Beckman Coulter, Inc., Pasadena, CA) was performed twice to ensure DNA fragments greater than 400 bp were removed completely. Adapters were diluted 1:20 to adjust for the low DNA concentration. Following adapter ligation (Figure B.1), additional cleanups were performed to remove adapter-dimers. The initial PCR followed the recommended protocol with a maximum of 12 cycles. Samples were divided into 5 ul aliquots and re-amplified until the final concentration of the pooled library was 100ng/ul. Subsequent amplifications were performed in a 25 ul reaction volume containing 5ul of the DNA library, 5x KAPA HiFi Fidelity Buffer, 0.3 mM each of the four deoxyribonucleoside 5'-triphosphates (dATP, dCTP, dGTP, and dTTP), 1 unit/ul A KAPA

HiFi DNA Polymerase (Kapa Biosystems, Wilmington, MA), 1x BSA, and 1 ul NextFlex primer oligonucleotide primer mix. PCR was run with an initial step of 95°C for4 min; with 12 cycles of 15 sec at 95°C; followed by 30 sec at 60°C, followed by 30 sec extension at 68°C; with a final hold at 4°C.

Each DNA sample was enriched for the complete mitogenome (Brandt *et al.* 2012) as well as three X-linked nuclear genes BGN, PLP, PHKA2 (Roca *et al.* 2005); three Toll-like receptor genes TLR3, TLR7, TLR8 (Astakhova *et al.* 2009); and the major histocompatibility complex gene DQA (Archie *et al.* 2010). In total ca. 36 kb was targeted for enrichment from each individual. DNA probe preparation (Figure B.1 - in box) followed the protocol outlined by Mason and colleagues (2011) with several modifications. Using DNA from a captive African savanna elephant each gene was amplified following the respective PCR protocols previously described (Archie *et al.* 2010; Astakhova *et al.* 2009; Brandt *et al.* 2012; Roca *et al.* 2005). Amplicons were grouped according to fragment size and pooled to 1 ug. Biotin-labeled probes were produced using the BioNick DNA Labeling System (Life Technologies; Grand Island, NY) following the recommended protocol. To ensure probes are within the optimal size range of 200-500 bp, agarose gel size checks were incorporated to empirically test the reaction conditions and terminate once the optimal size range was achieved. Probe concentration was checked by flourometric quantitation using the Qubit dsDNA HS Assay Kit (Life Technologies; Grand Island, NY). Probes were pooled according to fragment size (so that each nucleotide was equally represented) and concentration to 100 ng for capture hybridization.

Capture hybridization and selection (Figure B.1 - bottom panel) followed the protocol described by Del Mastro and Lovette (Del Mastro & Lovett 1997) with minor modifications described by Mason and colleagues (Mason *et al.* 2011). Hybridization reaction mix included adapter ligated DNA, biotin-labeled probe, hybridization buffer and mineral oil overlay. Samples were denatured at 99^{o}C for 5 minutes then incubated for 50 hours at 65-60 ^{o}C, decreasing about 2 ^{o}C every 24 hours. This procedure increases the retrieval of more divergent DNA sequences relative to the probe. Once hybridized, the reaction mixture was added to magnetic streptavidin coupled beads. Following several washes and an elution step, enriched DNA libraries will be amplified by PCR and sequenced on the Illumina MiSeq V2 platform.

Sequence reads were trimmed for adapters and quality using the "fastq-mcf" program from the ea-utils package (http://code.google.com/p/ea-utils) and assembled using Bowtie 2

(Langmead *et al*. 2012) or MITO-bim (Hahn *et al.* 2013) using an African savanna elephant reference sequences for each locus obtained from Genbank. Assembled sequences were visualized using Sequencher 5.1 and the next-generation sequencing tool-kit (Gene Codes Corporation, Ann Arbor, MI). Consensus sequences were verified as correct using a blast search.

Results

At the time of this writing, the methods described here are still being optimized. Initial testing has used 12 dung samples from wild African elephants collected in the Ziama Forest Nature Reserve in Guinea-Conakry, and Chobe National Park in Botswana. Negative controls were included for all PCR amplifications. The major hindrance to progress is a failure of the size selection step to remove fragments greater than 500 bp. To overcome this, the size selection step was repeated; though, success or failure has not yet been determined. Alternatively the gel method for size selection may be tried with the caveat that the gel-free (bead) method should retain more DNA and the gel method while more specific may result in a greater loss of already low DNA amounts.

Initial testing of the biotinylated bait and hybridization was apparently successful. Though tried on the DNA library with fragments greater than 500 bp, PCR amplification post enrichment yielded appropriately sized bands (albeit with streaks extending greater than 500 bp). Sequencing was not attempted as the banding pattern greater than 500 bp was considered to be sub-optimal. However, if further size selections attempts continue to be unsuccessful then sequencing should be attempted despite the sub-optimal quality.

Potential Benefits of this Research

Mitochondrial DNA has been used in numerous conservation genetic studies; specifically the control region, which has been used extensively to study population structure (Lu et al. 2001; Mock et al. 2002; Zhong et al. 2013). The control region is non-coding and thus highly variable due to greater mutation rates compared with that of coding regions (mutations in the control

region are neutral and are therefore not under selective pressures). Increased variation in a short evolutionary time is useful for distinguishing recent divergence but in older lineages the control region is more likely to become saturated, and it does not conform to a molecular clock hypothesis (Ingman et al. 2000). Mitochondrial DNA has been useful in studying African elephant phylogenetics (Brandt et al. 2012) and population structure (Georgiadis et al. 1994). Recently mitochondrial DNA sequences were examined from African elephants continent wide and the regional distribution of haplotypes was established (Ishida et al. 2013). The results from this study support using larger mitochondrial DNA sequences when determining the source of confiscated ivory. Ishida and colleagues (2013) showed that 62% of the haplotypes determined from 316 bp of the control region were found in a single country, versus the 84% of haplotypes identified from 4258 bp spanning the ND5 gene and part of the control region.

Establishing the provenance of illegal ivory is most successful when both mitochondrial and nuclear loci are utilized (Ishida et al. 2013). Nuclear genes are not subject to the same inheritance limitations as mitochondrial DNA because nuclear markers are dispersed by male elephants. X-linked genes BGN, PLP and PHKA2 have been typed in a large sample of African elephants and support a two-species model, emphasizing restricted gene flow between forest and savanna elephants (Roca et al. 2005; Roca et al. 2001). Recently species diagnostic single nucleotide polymorphisms have been detected among these three genes that are able to differentiate among all three extant elephant species: African savanna elephant, African forest elephant and Asian elephant (Ishida et al. 2011a). While nuclear and mitochondrial DNA produce incongruent phylogenetic patterns, nuclear markers are better indicators of gene flow and population connectivity in African elephants.

Traditionally microsatellites or short tandem repeats (STRs) have been used to infer recent population changes, but these are neutral markers from which selective pressures cannot be inferred. Immune genes are an effective indicator of selection on functionally important loci. Toll-like receptors are a unique class of genes that code for transmembrane proteins integral in immune response and have been described in numerous species (Cowled et al. 2011; Schneberger et al. 2011). A recent study has shown that these genes can be useful in determining genetic diversity, as well as to elucidate the dynamics between natural selection (balancing or directional) and genetic drift in populations of threatened species (Grueber et al. 2012). In elephants, primers have already been developed for Toll-like receptor genes 3, 7 and 8; though

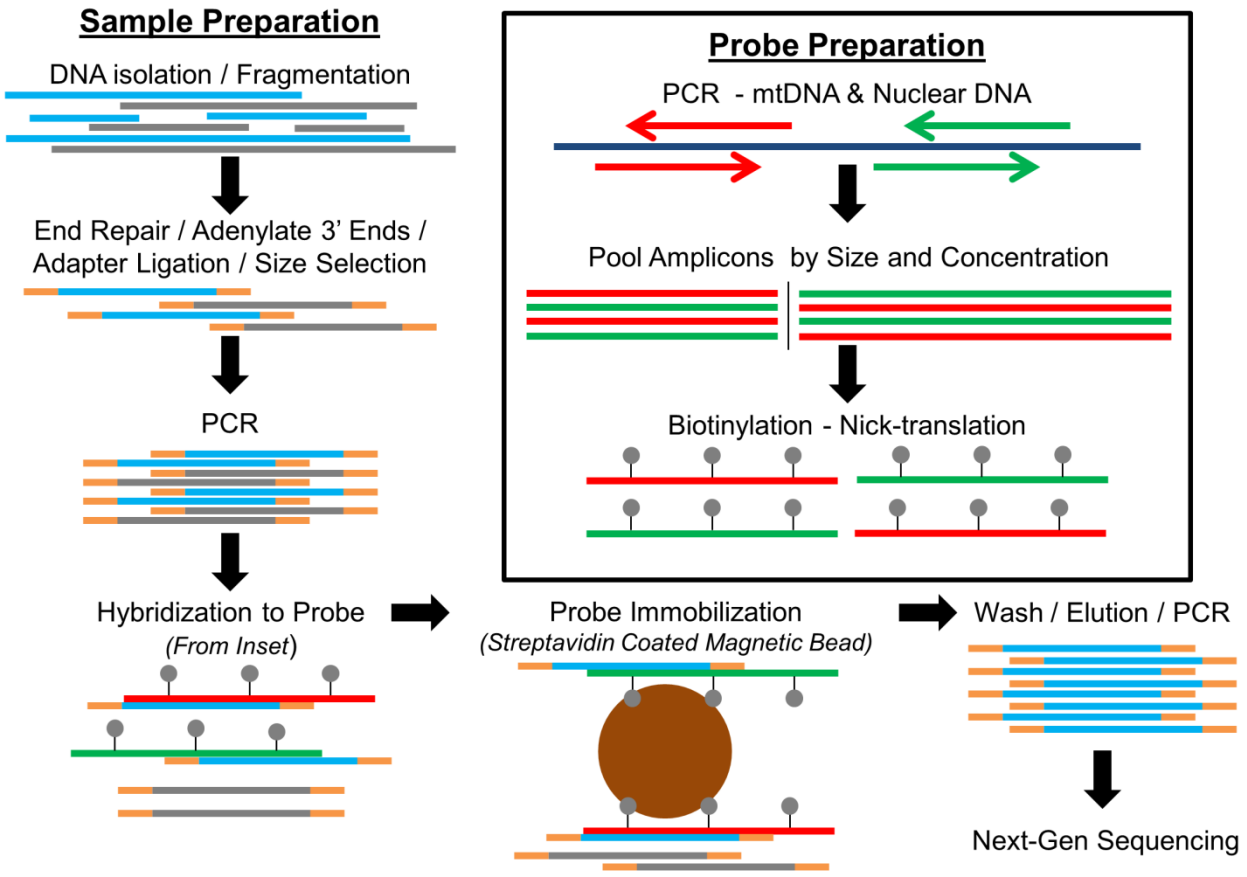
their only application has been to compare the phylogenetic relationship among other vertebrates (Astakhova et al. 2009). The major histocompatibility (MHC) class II DQA gene complex is one of the most polymorphic gene sets and is found in all jawed vertebrates (Ujvari & Belov 2011). MHC genes are integral in a number of immune functions and have been shown to provide an indirect measure of population fitness. In conservation studies, these genes have been shown to be useful for determining genetic diversity, disease susceptibility, identifying conservation units, and understanding local adaptions (Ujvari & Belov 2011). In African and Asian elephants the MHC complex was described (Archie et al. 2010); this locus was highly variable among 18 African elephants, with six alleles and a mean heterozygosity of 0.611.

Initially the protocol described here is being developed and tested using DNA from representative dung and/or ivory samples. Once successful, the techniques will be applied to a much larger sample of ca. 800 elephants. Additionally, dung samples are continually being collected from unstudied elephant populations. Using full mitochondrial genomes and large nuclear genes (c.a. 500 – 3000 bp), population structure and genetic diversity will be examined to better understand African elephant genetics.

Figures and Tables

Figure B.1. Overview of enrichment and capture hybridization method.
Sample preparation begins with DNA isolation followed by fragmentation of target (blue) and non-target (gray) DNA. Library preparation is performed following the manufacturer's recommended protocol with modifications described in Materials and Methods. Library is amplified by PCR. Probe preparation is described in the box. Loci of interest are amplified by PCR from a high quality DNA source from *Loxodonta africana*. Products are pooled by amplicon size in equal concentrations. Products are biotinylated by nick-translation following the manufacturer's recommended protocol with modifications described in Materials and Methods. Biotinylated products are again pooled so that each nucleotide in the probe is equally represented. Each library is enriched by hybridization to the probe and immobilized using streptavidin coated magnetic beads. Non-target DNA (gray) is removed while target DNA (blue) is purified. The library now containing primarily enriched target DNA is amplified by PCR and sequenced by next-generation sequencing technology.

Figure B.1. Continued.



Literature Cited

Amos W, Whitehead H, Ferrari MJ, *et al.* (1992) Restrictable dna from sloughed cetacean skin - its potential for use in population analysis. *Marine Mammal Science* **8**, 275-283.

Archie EA, Henry T, Maldonado JE, *et al.* (2010) Major histocompatibility complex variation and evolution at a single, expressed DQA locus in two genera of elephants. *Immunogenetics* **62**, 85-100.

Archie EA, Hollister-Smith JA, Poole JH, *et al.* (2007) Behavioural inbreeding avoidance in wild African elephants. *Molecular Ecology* **16**, 4138-4148.

Astakhova NM, Perelygin AA, Zharkikh AA, *et al.* (2009) Characterization of equine and other vertebrate TLR3, TLR7, and TLR8 genes. *Immunogenetics* **61**, 529-539.

Barnes RFW (1999) Is there a future for elephants in West Africa? *Mammal Review* **29**, 175-199.

Barriel V, Thuet E & Tassy P (1999) Molecular phylogeny of Elephantidae. Extreme divergence of the extant forest African elephant. *C R Acad Sci III* **322**, 447-454.

Blanc JJ, Barnes RFW, Craig GC, *et al.* (2007) African elephant status report 2007. An update from the African elephant database. *Occasional Papers of the IUCN Species Survival Commission* **33**, i-vi, 1-275.

Brandt AL, Ishida Y, Georgiadis NJ & Roca AL (2012) Forest elephant mitochondrial genomes reveal that elephantid diversification in Africa tracked climate transitions. *Molecular Ecology* **21**, 1175-1189.

Cahill JA, Green RE, Fulton TL, *et al.* (2013) Genomic Evidence for Island Population Conversion Resolves Conflicting Theories of Polar Bear Evolution. *PLoS Genetics* **9**, 1-8.

Comstock KE, Georgiadis N, Pecon-Slattery J, *et al.* (2002) Patterns of molecular genetic variation among African elephant populations. *Mol Ecol* **11**, 2489-2498.

Cui Y, Lindo J, Hughes CE, *et al.* (2013) Ancient DNA analysis of mid-Holocene individuals from the northwest coast of North America reveals different evolutionary paths for mitogenomes. *Plos One* **8**.

Debruyne R (2005) A case study of apparent conflict between molecular phylogenies: the interrelationships of African elephants. *Cladistics* **21**, 31-50.

Del Mastro RG & Lovett M (1997) Isolation of coding sequences from genomic regions using direct selection. *Methods in molecular biology (Clifton, N.J.)* **68**, 183-199.

Douglas-Hamilton I (1987) African elephants: population trends and their causes. *Oryx* **21**, 11-24.

Eggert LS, Patterson G & Maldonado JE (2008) The Knysna elephants: a population study conducted using faecal DNA. *African Journal of Ecology* **46**, 19-23.

Eggert LS, Rasner CA & Woodruff DS (2002) The evolution and phylogeography of the African elephant inferred from mitochondrial DNA sequence and nuclear microsatellite markers. *Proc R Soc Lond B Biol Sci* **269**, 1993-2006.

Georgiadis N, Bischof L, Templeton A, *et al.* (1994) Structure and history of African elephant populations: I. Eastern and southern Africa. *J Heredity* **85**, 100-104.

Groves CP & Grubb P (2000) Do *Loxodonta cyclotis* and *L. africana* interbreed? *Elephant* **2**, 4-7.

Grubb P, Groves CP, Dudley JP & Shoshani J (2000) Living African elephants belong to two species: *Loxodonta africana* (Blumenbach, 1797) and *Loxodonta cyclotis* (Matschie, 1900). *Elephant* **2**, 1-4.

Hahn C, Bachmann L & Chevreux B (2013) Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads-a baiting and iterative mapping approach. *Nucleic Acids Research* **41**.

Hollister-Smith JA, Poole JH, Archie EA, *et al.* (2007) Age, musth and paternity success in wild male African elephants, *Loxodonta africana*. *Animal Behaviour* **74**, 287-296.

Ishida Y, Demeke Y, de Groot PJV, *et al.* (2011a) Distinguishing forest and savanna African elephants using short nuclear DNA sequences. *Journal of Heredity* **102**, 610-616.

Ishida Y, Georgiadis NJ, Hondo T & Roca AL (2013) Triangulating the provenance of African elephants using mitochondrial DNA. *Evolutionary Applications* **6**, 253-265.

Ishida Y, Oleksyk TK, Georgiadis NJ, *et al.* (2011b) Reconciling apparent conflicts between mitochondrial and nuclear phylogenies in African elephants. *Plos One* **6**, 16.

Johnson MB, Clifford SL, Goossens B, *et al.* (2007) Complex phylogeographic history of central African forest elephants and its implications for taxonomy. *BMC Evol Biol* **7**, 244.

Mailand C & Wasser SK (2007) Isolation of DNA from small amounts of elephant ivory. *Nat Protoc* **2**, 2228-2232.

Mason VC, Li G, Helgen KM & Murphy WJ (2011) Efficient cross-species capture hybridization and next-generation sequencing of mitochondrial genomes from noninvasively sampled museum specimens. *Genome Research* **21**, 1695-1704.

Perry GH, Marioni JC, Melsted P & Gilad Y (2010) Genomic-scale capture and sequencing of endogenous DNA from feces. *Molecular Ecology* **19**, 5332-5344.

Petit RJ & Excoffier L (2009) Gene flow and species delimitation. *Trends Ecol Evol* **24**, 386-393.

Poole JH (1989a) Mate guarding, reproductive success and female choice in African elephants. *Animal Behaviour* **37**, 842-849.

Poole JH (1989b) Announcing intent: the aggressive state of musth in African elephants. *Animal Behaviour* **37**, 140-152.

Roca AL, Georgiadis N & O'Brien SJ (2005) Cytonuclear genomic dissociation in African elephant species. *Nat Genet* **37**, 96-100.

Roca AL, Georgiadis N, Pecon-Slattery J & O'Brien SJ (2001) Genetic evidence for two species of elephant in Africa. *Science* **293**, 1473-1477.

Rohland N, Reich D, Mallick S, *et al.* (2010) Genomic DNA sequences from mastodon and woolly mammoth reveal deep speciation of forest and savanna elephants. *PLoS Biol* **8**, e1000564.

Species Survival Network (2007) Inconvenient but true: The unrelenting global trade in elephant ivory. Species Survival Network, The Hague.

Wasser SK, Clark B & Laurie C (2009) The ivory trail. *Sci Am* **301**, 68-74, 76.

Wasser SK, Joseph Clark W, Drori O, *et al.* (2008) Combating the illegal trade in African elephant ivory with DNA forensics. *Conserv Biol* **22**, 1065-1071.

Wasser SK, Mailand C, Booth R, *et al.* (2007) Using DNA to track the origin of the largest ivory seizure since the 1989 trade ban. *Proc Natl Acad Sci U S A* **104**, 4228-4233.

Wasser SK, Shedlock AM, Comstock K, *et al.* (2004) Assigning African elephant DNA to geographic region of origin: applications to the ivory trade. *Proc Natl Acad Sci U S A* **101**, 14847-14852.

APPENDIX C: PROTOCOL FOR DNA ENRICHMENT FROM LOW QUALITY SAMPLES BY CAPTURE HYBRIDIZATION

I. Sample DNA preparation

1. DNA isolation
    a. Follow protocol for Qiagen Mini Stool Kit (dung)
    b. Ivory protocol – Mailand and Wasser 2007 Nature protocols
2. Quality check – check DNA concentration with Qubit HS to be sure isolation worked

II. Library Prep (NEXTflex DNA Sequencing Kit w/ Mods – Cui et al. 2013 PLoS One)

1. DNA fragmentation
    a. Sonicate to < 800 bp
        - microTUBE AFA Fiber Snap-Cap – 130 µl sample
        - Target BP (peak) = 800
        - Peak incident power (W) = 50
        - Duty factor = 5%
        - Cycles per burst = 200
        - Treatment time (s) = 80
        - Temperature (^{o}C) = 20
        - Sample volume (ul) = 130
        - http://covarisinc.com/wp-content/uploads/pn_010166.pdf
        - **Can be done with 50ul screw cap, just follow appropriate settings
2. End Repair
    a. As per manual
3. Gel-free size selection Clean-up (B1)
    a. As per manual
    b. Repeat for a second size selection to ensure fragments >400bp removed
4. 3' Adenylation
    a. As per manual
5. Adapter Ligation
    a. 1:20 dilution of adapters/barcode - DNA concentration in extract is low
    b. Remaining steps as per manual
6. Clean-up
    a. As per manual
    b. Repeat for a second clean up to remove adapter-adapter ligation
7. PCR amplification of genomic library
    a. Initial amplification – as per manual (maximum 12 cycles)
    b. Purification
        - Zymo DNA Clean & Concentrator-5 kit
        - Elute to 50ul
        - Divide into 5ul aliquots (may lose some volume from spin column)
    c. Second amplification
        - 25 ul reaction volume
            1. Library (aliquot from step 7b) ..... 5 ul
            2. H_2O ..... 11.75 ul

3. 5x KAPA Buffer ..... 5 ul
4. 10mM dNTP mix ..... 0.75 ul
5. NextFlex primer mix ..... 1 ul
6. BSA ..... 1 ul
7. KAPA polymerase ..... 0.5 ul

- Thermoprofile
  1. 95°C – 4 min
  2. 12 cycles
     a. 95°C – 15s
     b. 60°C – 30s
     c. 68°C – 30s
  3. 4°C – hold

d. Repeat step C and pool secondary amplifications until concentration is 100 ng/ul
   - Quantify with Qubit HS

e. Purification
   - Zymo DNA Clean & Concentrator-5 kit
   - Elute so that concentration is 500ng to 1ug in < 15 ul total volume

III. Capture probe generation (Mason et al. 2011 – Genome Research)

1. PCR amplification using zoo LAF
   a. MtDNA (8 sets)
   b. BGN (1 set)
   c. PLP (1 set)
   d. PHK (1 set)
   e. TLR 3 (2 sets)
   f. TLR 7 (2 sets)
   g. TLR 8 (1 set)
   h. MHC-DQA (3 sets)
2. Pool each replicate, purify using Zymo DNA Clean & Concentrator-5 kit, and quantify
   a. Quantify PCR product, if greater than 5ug use QIAquick PCR Purification Kit
3. Pool PCR products (based on concentration) to 1ug
   a. MIN (6) ..... 0.1667 ug each
   b. Q2 (2) ..... 0.5 ug each
   c. MED (5) ..... 0.2 ug each
   d. Q3 (5) ..... 0.2 ug each
   e. MAX (2) ..... 0.5 ug each
4. Labeling (BioNick DNA Labeling System)
   a. Follow manufacturer protocol – adjust for smaller starting fragment (1/2 time)
   b. See Biotin-High Prime protocol for size check before stopping procedure
   c. Quatify
   d. Pool biotin-labeled probes based on average fragment size to total 100ng
      - MIN (6) ..... 11.6 ng
      - Q2 (2) ..... 5.7 ng
      - MED (5) ..... 28.4 ng
      - Q3 (5) ..... 36.4 ng
      - MAX (2) ..... 17.9 ng

IV. Capture hybridization and selection (Del Mastro and Lovett 1997 – Genome Res)

1. Hybridization
    a. Reaction mix
        - 500 ng to 1 ug ..... amplified, adapter ligated DNA
        - 100 ng ..... biotin-labeled mtDNA probe
        - Equal volume ..... 2x hybridization buffer (not to exceed 15ul)
        - 50 ul (overlay) ..... mineral oil
    b. Denature for 5 minutes at 99 ^{o}C
    c. Incubate for 50 hours at 65-60 ^{o}C (reduce ~2 ^{o}C every 24hrs)
2. Capture
    a. Prepare beads
        - 1mg ..... Dynabeads M-280 Streotavidin beads (Dynnal)
            - Wash 3 times with 100 ul TEN buffer in magnetic tube holder
            - Re-suspend in 100 ul TEN buffer
    b. Washes
        i. Remove bottom layer of hybridization reaction mix away from the oil
        ii. Add hybridization mix to dynabeads, mix gently – incubate at room temp for 20 minutes (mix occasionally to prevent beads from settling)
        iii. Place in magnetic tube holder for 1 minute
        iv. Remove and retain supernatant
            - Quality check - *Keep the supernatant and the wash solutions just in case the direct selection failed because of faulty reagents. If this occurs, take 1 ul from the supernatant and the wash solutions, and PCR-amplify them Evaluate the PCR products by electrophoresis on a 1% agarose gel. Southern blot or vacuum blot the gel onto a nylon membrane and hybridize the positive reporter cDNA to the filter to determine if any of the lanes show a smear The result may indicate that either the cDNA did not hybridize to the template or that the cDNA was washed off the beads during the wash steps. If either of these appears to be the case, it is advisable to make fresh solutions and start again*
        v. Remove tube from magnetic tube holder, re-suspend beads in 1 ml of wash solution 1 (low stringency)
        vi. Incubate at room temperature for 15 minutes
        vii. Return tube to magnetic tube holder for 1 minute
        viii. Remove supernatant
        ix. Repeat steps v-viii
        x. Repeat steps v-vii 3 times with wash solution 2, incubating at 65 ^{o}C for 15 minutes (instead of room temp)

c. Elution
   - 25 ul ..... 0.1 N NaOH
     - Add to beads
     - Incubate at room temperature for 20 minutes
     - Gently vortex every 5 minutes (prevent beads from settling)
   - 25 ul ..... Tris-HCl (pH 7.5)
     - Neutralize reaction
   - Purify with CentriSpin20 column (Princeton Separations)

V. Post-enrichment PCR

1. 25 ul reaction volume
   a. Library (aliquot from step 7b) ..... 5 ul
   b. H_2O ..... 11.75 ul
   c. 5x KAPA Buffer ..... 5 ul
   d. 10mM dNTP mix ..... 0.75 ul
   e. NextFlex primer mix ..... 1 ul
   f. BSA ..... 1 ul
   g. KAPA polymerase ..... 0.5 ul
2. Thermoprofile
   f. 95^oC – 4 min
   g. 12 cycles
      - 95^oC – 15s
      - 60^oC – 30s
      - 68^oC – 30s
   h. 4^oC – hold

1. Purification
   - Zymo DNA Clean & Concentrator-5 kit

2. Resolve on 1% Agarose gel

VI. Next-generation sequencing

1. Verify quality - Agilent bioanalyzer chip
2. Submit to core – contact for concentrations / multiplexing / pooling

VII. Analyze

1. Assembly with Bowtie or Mito-Bim
2. Sequencher w/ NGS toolkit

Buffers and Solutions

- 2x Hybridization Buffer — 50 ml total volume
  - 1.5mM NaCl — 4.383 mg
  - 40 mM sodium phosphate* — 2 ml (1M stock)
  - 10 mM EDTA — 146.12 mg
  - 10x Denhardt's solution — 10 ml (50x stock)
  - 0.2% sodium dodecyl sulfate (SDS) — 100 mg

  ** 1M Sodium phosphate buffer (50 ml)*
    - *0.183M NaH_2PO_4(Mono)* — *1.2627 g*
    - *0.778M Na_2HPO_4(Di)* — *5.5238 g*
    - *H_2O* — *bring to 50 mL*

- TEN Buffer — 50 ml total volume
  - 10 mM Tris-HCl, pH 7.5* — 0.5 ml (1M stock)
  - 1 mM EDTA — 14.612 mg
  - 1 M NaCl — 2.922 g

  * 1M Tris-HCl
    - *Tris-base* — *6.057 g*
    - *H_2O* — *40 ml*
    - *HCl* — *~3.25 ml (adjust pH to 7.5)*
    - *H_2O* — *bring to 50 mL*

- Wash solution I - low stringency — 50 ml total volume
  - 1X saline-sodium citrate (SSC) — 2.5 ml (20x stock)
  - 0.1% sodium dodecyl sulfate (SDS) — 50 mg

- Wash solution II - high stringency — 50 ml total volume
  - 0.1X saline-sodium citrate (SSC) — 250 ul (20x stock)
  - 0.1% sodium dodecyl sulfate (SDS) — 50 mg